МАЈ - ЈУЛИ, 2022.

ИНФОРМАТИЧКА ТЕОРИЈА

ПРИЛОЗИ ПРИВАТНИМ РАЗГОВОРИМА

РАСТКО ВУКОВИЋ
ПРОФ. МАТЕМАТИКЕ
ГИМНАЗИЈА БАЊА ЛУКА

Растко Вуковић, мај-јули 2022.
Информатичка Теорија – прилози приватним разговорима
https://archive.org/details/it_20220709/

Предговор

Прилози су нека од мојих приватних израчунавања која сам обављао припремајући одговоре на (замишљена или стварна) питања ми о (углавном мојој) „теорији информације“, а која сам потом организовао у Блогу (http://rvukovic.net/blog/). Аутор, маја 2022.

Почетак јуна је 2022. и тек сам стигао до наслова „29. Закон одржања“, а умишљао сам да ће списи до данас бити завршени. Није ово кашњење због моје лењости (не кажем да је нема), нити је оно у (иначе споријем) откривању нових закона, јер за сада пишем по сећању, са папирића, примедби и описа из дописивања, него су разлози у обиму који ме је изненадио. До краја јуна стигао сам да сакупим текстова до наслова „41. Пренос канала“.

Последњи наслов овде „43. Условне вероватноће“ писао сам почетком јула 2022. године. Ма како се трудио да згуснем те формуле и кратка објашњења, испоставило се да се претходних година тога нагомилало превише за само једну скрипту и да ће јој, овој, можда уследити наставак.

Аутор, 2022.

Садржај

I Део: Канали

Линеарне операторе, њихове матрице, системе линеаних једначина, као и све врсте преносника информација овде једноставно називамо „каналима“. Тиме остављамо могућност да се пренос података може одвијати и нелинеарним каналима, којима се за сада нећемо бавити.

На много места наћи ћете систематски развој познатих појмова линеарне алгебре, па и у мојим текстовима, од скрипти за гимназијску наставу до квантне механике или теорије информације (Простор-Време). Зато те делове овде прескачем. Били би губљење времена за читаоце који те области математике познају и сада би хтели видети нешто ново.

Ако сте навикли читати уводе у алгебру, наизглед би вам ова могла изгледати збрда-здола писана, а заправо је уредно послагана како бих се могао на делове позивати у одговорима на евентуална ми питања. Не би ту требало бити претешких ствари – за макар каквог познаваоца линеарне алгебре – јер ту су да такве појасне. Као што је једном приликом Ајнштајн наводно рекао, ако не умеш нешто објаснити једноставно, онда то можда и не разумеш. Не знам колико ћете наћи да сам у томе успео, али избегавао сам оно што нисам разумео.

Елем, за теорију информације за коју су ови текстови, све је у наводним каналима преноса података. Или је то тако или се само правимо да тако бива док не утврдимо супротно. Зато ова прича ризикује да буде преопширна, јер се не надам контрадикцијама. Такође, јер је информација ткиво простора, времена и материје, а неизвесност њена суштина, па је грађевина ових канала на вероватноћама. Познате теорије о њима ће нам такође требати, а посебно расподеле, јер ни око таквих нећемо губити време.

Дакле, не подучавамо, већ продубљујемо знања. Сигуран сам да до 2020. нисте приметили да заменом „тачно“ са „нетачно“ и обрнуто у формулама алгебре логике добијамо еквивалентне системе „света лажи“ и „света истина“, на начин да су контрадикција првог и таутологја другог замењиве варијабле. Или, да се тополошки може доказивати егзистенција додатних временских димензија, затим да из сконости природе чешћој реализацији вероватнијих исхода следи њена склоност развоју у мање информативна стања, па и принцип најмањег дејства.

Домети алгебре матрица сежу даље него што можемо замислити. Зато ће приче о њој као о каналима преноса података, чак и након оваквих, увек остајати посебне и важне.

01. Неквадратна матрица канала $\hat{K}$

Питање: На питање „Мора ли матрица пресликавања бити квадратна?“, одговарате одрично и наводите пример са семафором (Semaphore) са неквадратном матрицом канала и додајете да „нема инверзне матрице“. Можете ли ми овај додатак елаборирати?

Одговор: Могу, рецимо са мало алгебре. Нека је матрица канала нека $\hat{K} = [\![k_{ij}]\!]$, попут оне из поменутог прилога о семафору. Типа је 2×3, тако да је $\vec{q} = \hat{K}\vec{p}$, где је $\vec{p}$ улазни 3-компонентни вектор, а $\vec{q}$ излазни 2-компонентни. Узмимо да јо је „инверзна“ матрица $\hat{K}' = [\![k'_{ij}]\!]$, типа 3×2, односно $\vec{p} = \hat{K}'\vec{q}$, или детаљније:

$$\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \end{bmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}, \quad \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{bmatrix} k'_{11} & k'_{12} \\ k'_{21} & k'_{22} \\ k'_{31} & k'_{32} \end{bmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}.$$

Сменом прве у другу једначину, добијамо $\vec{p} = \hat{K}'\vec{q} = \hat{K}'(\hat{K}\vec{p}) = (\hat{K}'\hat{K})\vec{p} = \hat{A}\vec{p}$, где је $\hat{A} = [\![a_{ij}]\!]$ квадратна матрица, реда 3. Дакле $\hat{K}'\hat{K} = \hat{A}$, или у развијеном облику:

$$\begin{bmatrix} k'_{11} & k'_{12} \\ k'_{21} & k'_{22} \\ k'_{31} & k'_{32} \end{bmatrix} \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Претпоставимо ли да је матрица $\hat{A}$ јединична, да на дијагонали има јединице ($a_{ii} = 1$), а да су сви остали њени елементи нуле ($a_{ij} = 0$, када $i \neq j$), онда њен први ред даје систем једначина:

$$\begin{cases} k'_{11}k_{11} + k'_{12}k_{21} = 1 = a_{11} \\ k'_{11}k_{12} + k'_{12}k_{22} = 0 = a_{12} \\ k'_{11}k_{13} + k'_{12}k_{23} = 0 = a_{13} \end{cases}$$

а аналогно други и трећи ред. Коефицијенти k'_{11} и k'_{12} матрице $\hat{K}'$ нису нуле, ако има било каквог преноса података. Детерминанта система доње две од ове три једначине је нула, да би постојало и нетривијално решење (различито од нуле). Отуда $k_{12}\colon k_{13} = k_{22}\colon k_{23}$. Поновимо ли исто за други и трећи ред, налазимо: $k_{11}\colon k_{13} = k_{21}\colon k_{23}$ и $k_{11}\colon k_{12} = k_{21}\colon k_{22}$, те је $k_{11}\colon k_{12}\colon k_{13} = k_{21}\colon k_{22}\colon k_{23}$. То значи да су коефицијенти првог и другог реда $\hat{K}$ пропорционални и да је систем у контрадикцији. Наиме, није могућа јединица у првој једначини ($a_{11} = 1$) а нуле у остале две ($a_{12} = a_{13} = 0$).

Претпоставимо даље да матрица $\hat{A}$ није јединична, него да има и ван дијагоналне коефицијенте различите од нуле (бар неки $a_{ij} \neq 0$, када $i \neq j$). Претходи систем посматрајмо тако да је $x = k'_{11}$, $y = k'_{12}$ и $z = 1$, па је десно од једнакости $a_{1j}z$, редом за $j = 1, 2, 3$. Пребацимо коефицијенте a_{1j} са десне на леву страну једнакости. Добили смо хомоген систем:

$$\begin{cases} k_{11}x + k_{12}y - a_{11}z = 0 \\ k_{12}x + k_{22}y - a_{12}z = 0 \\ k_{13}x + k_{23}y - a_{13}z = 0 \end{cases}$$

а аналогно за други и трећи ред множења претходних матрица. Да би овај систем имао нетривијално решење (различито од нуле), а иначе већ је $z = 1$, он не може бити регуларан. Његова детерминанта мора бити нула. Отуда:

$$\begin{vmatrix} k_{11} & k_{12} & a_{11} \\ k_{12} & k_{22} & a_{12} \\ k_{13} & k_{23} & a_{13} \end{vmatrix} = 0,$$

$$a_{11}\begin{vmatrix} k_{12} & k_{22} \\ k_{13} & k_{23} \end{vmatrix} + a_{12}\begin{vmatrix} k_{13} & k_{23} \\ k_{11} & k_{12} \end{vmatrix} + a_{13}\begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} = 0.$$

Урадимо аналогно за други и трећи ред множења матрица ($\hat{K}'\hat{K} = \hat{A}$) и уведимо скраћене ознаке за ове детерминанте другог реда. Формирамо систем, по непознатима тим детерминантама:

$$\begin{cases} a_{11}D_1 + a_{12}D_2 + a_{13}D_3 = 0 \\ a_{21}D_1 + a_{22}D_2 + a_{23}D_3 = 0 \\ a_{31}D_1 + a_{32}D_2 + a_{33}D_3 = 0 \end{cases}$$

Тај је систем хомоген и не може имати само једно, тада нулто решење ($D_1 = D_2 = D_3 = 0$), јер би то значило да по горњим матрицама нема било каквог преноса. Дакле, мора бити:

$$\det \hat{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0$$

а то је контрадикција са претпоставком ($\hat{K}'\hat{K} = \hat{A}$) да матрице канала нису нулте, да бар нешто преносе.

Нема могућности да матрица $\hat{A} = \hat{K}'\hat{K}$ буде макар нека блага „инверзна“, да буде било шта осим нула-матрице. Толико о овом редоследу множења – када се добија матрица 3×3, када је више једначина (9) од непознатих (6).

Сасвим је другачија ситуација након множења другим редоследом, када се добија матрица 2×2, а број непознатих (6) је већи од броја једначина (4). Јасно је да тада очекујемо појаву решења и то не само једног.

Дакле, претпостављену матрицу $\hat{K}'$ и формулу $\vec{p} = \hat{K}'\vec{q}$ сада увршtавамо у $\vec{q} = \hat{K}\vec{p}$, па добијамо:

$$\vec{q} = \hat{K}\vec{p} = \hat{K}(\hat{K}'\vec{q}) = (\hat{K}\hat{K}')\vec{q} = \hat{A}'\vec{q}.$$

Отуда:

$$\hat{A}' = \hat{K}\hat{K}' = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \end{bmatrix} \begin{bmatrix} k'_{11} & k'_{12} \\ k'_{21} & k'_{22} \\ k'_{31} & k'_{32} \end{bmatrix} =$$

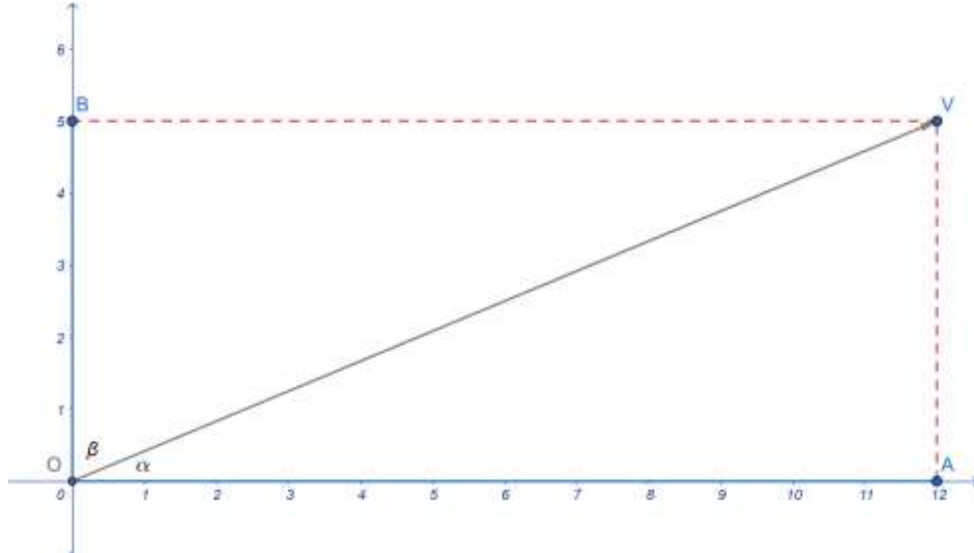
$$= \begin{bmatrix} k_{11}k'_{11} + k_{12}k'_{21} + k_{13}k'_{31} & k_{11}k'_{12} + k_{12}k'_{22} + k_{13}k'_{32} \\ k_{21}k'_{11} + k_{22}k'_{21} + k_{23}k'_{31} & k_{21}k'_{12} + k_{22}k'_{22} + k_{23}k'_{32} \end{bmatrix}$$

$$= \begin{bmatrix} \vec{k}_1 \cdot \vec{k'}_1 & \vec{k}_1 \cdot \vec{k'}_2 \\ \vec{k}_2 \cdot \vec{k'}_1 & \vec{k}_2 \cdot \vec{k'}_2 \end{bmatrix},$$

где је стављено $\vec{k}_i = (k_{i1}, k_{i2}, k_{i3})$, $\vec{x}, y = \vec{k'}_j = (k'_{1j}, k'_{2j}, k'_{3j})$, за $i, j = 1, 2$. А то је још једна занимљива ситуација коју треба пажљиво анализирати, неком другом приликом.

02. Координатна вероватноћа

Посматрајмо вектор $\vec{v} = \overrightarrow{OV}$ у Декартовом правоуглом систему координата OXY, на слици.



Углови које вектор $\vec{v}$ гради са апсцисом (x-осом) и ординатом (y-осом) су α и β, па су пројекције на њих $\overline{OA} = \overline{OV}\cos\alpha$ и $\overline{OB} = \overline{OV}\cos\beta$. Ови углови су комплементни ($\alpha + \beta = 90°$) и на основу тригонометријских идентитета добијамо $\cos^2\alpha + \cos^2\beta = 1$. Наиме, збир квадрата синуса и косинуса истог угла је један, а синус угла је косинус комплемента. Када је дати вектор јединични, интензитета $\overline{OV} = 1$, онда су његове (окомите) пројекције на координатне осе косинуси углова које вектор заклапа са осама, а збир квадрата пројекција је један.

Аналогно важи уопште за више димензионални Декартов правоугли систем ($Ox_1x_2 \dots x_n$). Ако су углови које јединични вектор, $|\vec{v}| = 1$, заклапа са осама $\alpha_k = \sphericalangle(x_k, \vec{v})$, редом за $k = 1, 2, \dots, n$, тада су пројекције тог вектора на осе косинуси ових углова, а збир квадрата тих косинуса је 1. Наиме, $\vec{v} = (v_1, v_2, \dots, v_k)$, а $v_k = \cos\alpha_k$, па квадрат дужине вектора $|\vec{v}|^2 = v_1^2 + v_2^2 + \dots + v_n^2 = 1$ постаје $\cos^2\alpha_1 + \cos^2\alpha_2 + \dots + \cos^2\alpha_n = 1$.

Репрезентацију ове координатне особине преносимо на вероватноће. Поменути јединични вектор, $\vec{v}$, представља случајни догађај са могућим исходима ($\omega_1, \omega_2, \dots, \omega_n$) које су приказане координатним осама. У случају расподеле вероватноћа, комплетног скупа дисјунктних исхода, поједине су вероватноће $P(\omega_k) = p_k$ квадрати косинуса углова између тог вектора и оса, дакле $p_k = \cos^2\alpha_k$ и $|\vec{v}|^2 = p_1 + p_2 + \dots + p_n = 1$.

Доследно редефинишемо и матрицу канала преноса порука (Channel):

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} \cos^2\alpha_1 \\ \cos^2\alpha_2 \\ \cos^2\alpha_3 \end{bmatrix} =$$

$$= \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \left(\begin{bmatrix} \cos\alpha_1 & 0 & 0 \\ 0 & \cos\alpha_2 & 0 \\ 0 & 0 & \cos\alpha_3 \end{bmatrix} \begin{bmatrix} \cos\alpha_1 \\ \cos\alpha_2 \\ \cos\alpha_3 \end{bmatrix} \right)$$

$$= \left(\begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} \cos\alpha_1 & 0 & 0 \\ 0 & \cos\alpha_2 & 0 \\ 0 & 0 & \cos\alpha_3 \end{bmatrix} \right) \begin{bmatrix} \cos\alpha_1 \\ \cos\alpha_2 \\ \cos\alpha_3 \end{bmatrix} =$$

$$= \begin{bmatrix} k_{11}\cos\alpha_1 & k_{12}\cos\alpha_2 & k_{13}\cos\alpha_3 \\ k_{21}\cos\alpha_1 & k_{22}\cos\alpha_2 & k_{23}\cos\alpha_3 \\ k_{31}\cos\alpha_1 & k_{32}\cos\alpha_2 & k_{33}\cos\alpha_3 \end{bmatrix} \begin{bmatrix} \cos\alpha_1 \\ \cos\alpha_2 \\ \cos\alpha_3 \end{bmatrix}$$

$$= \begin{bmatrix} k_{11}\cos^2\alpha_1 + k_{12}\cos^2\alpha_2 + k_{13}\cos^2\alpha_3 \\ k_{21}\cos^2\alpha_1 + k_{22}\cos^2\alpha_2 + k_{23}\cos^2\alpha_3 \\ k_{31}\cos^2\alpha_1 + k_{32}\cos^2\alpha_2 + k_{33}\cos^2\alpha_3 \end{bmatrix}.$$

Са друге стране и вектор q можемо посматрати као да је изведен из неког јединичног $\vec{u}$ чији углови према истим осама су β_1, β_2 и β_3, тако да је $|\vec{u}|^2 = q_1 + q_2 + \dots + q_n = 1$, а квадрати његових пројекција су вероватноће $q_k = \cos^2\beta_k$ ($k = 1, 2, \dots, n$). Али подсетимо се прво шта су то „дисјунктни догађаји“ и шта су „условне вероватноће“.

Дефиниција 1. Догађаји су дисјунктни ($AB = \emptyset$) када се не могу оба десити истовремено. Ако је A_k потпун скуп дисјунктних догађаја онда је пресек парова, или производ $A_i A_j = \emptyset$ када $i \neq j$, а унија или збир свих $\bigcup_k A_k = \Omega$. Тада $P(A_k) = p_k$ чини једну расподелу вероватноћа ($\sum_k p_k = 1$) које користимо у овим задацима. ▨

Дефиниција 2. Условна вероватноћа догађаја A ако се десио догађај B је $P(A|B) = P(AB)/P(B)$. На пример, из шпила од 52 карте, нека су догађаји A и B посебно да су случајно изабране карте „ас“ и „пик“. Производ (пресек) ових AB је догађај да је изабрана карта „пик ас“. Тада је $P(AB) = 1/52$, затим $P(B) = 13/52$, па је условна вероватноћа $P(A|B) = 1/13$, да је изабрана карта „ас“ ако знамо да је „пик“. Иначе је $P(A|B) \geq 0$, $P(\Omega|B) = 1$, где је Ω сигуран догађај – скуп свих могућности. ▨

Дефиниција 3. Догађаји A и B независни су када је $P(AB) = P(A)P(B)$. Произвољан и немогућ догађај су независни. Ако су догађаји независни и бар један од њих је немогућ, онда су они дисјунктни. Ако су догађаји дисјунктни, а ни један од њих није немогућ они су тада зависни. Када су A и B независни догађаји, онда су A и $\bar{B}$ независни, такође $\bar{A}$ и B, или $\bar{A}$ и $\bar{B}$. Догађај $\bar{B}$ је супротан, комплементаран са B, што значи да ће се први десити ако и само ако се други неће десити. ▨

Сада формирајмо пристигле сигнале на сличан начин како смо то урадили са послатим:

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} \cos^2\beta_1 \\ \cos^2\beta_2 \\ \cos^2\beta_2 \end{bmatrix} = \begin{bmatrix} \cos\beta_1 & 0 & 0 \\ 0 & \cos\beta_2 & 0 \\ 0 & 0 & \cos\beta_3 \end{bmatrix} \begin{bmatrix} \cos\beta_1 \\ \cos\beta_2 \\ \cos\beta_3 \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\beta_1 & 0 & 0 \\ 0 & \cos\beta_2 & 0 \\ 0 & 0 & \cos\beta_3 \end{bmatrix} \begin{bmatrix} \cos\beta_1(\cos^2\alpha_1 + \cos^2\alpha_2 + \cos^2\alpha_3) \\ \cos\beta_2(\cos^2\alpha_1 + \cos^2\alpha_2 + \cos^2\alpha_3) \\ \cos\beta_3(\cos^2\alpha_1 + \cos^2\alpha_2 + \cos^2\alpha_3) \end{bmatrix}$$

$$= \begin{bmatrix} \cos\beta_1 & 0 & 0 \\ 0 & \cos\beta_2 & 0 \\ 0 & 0 & \cos\beta_3 \end{bmatrix} \left(\begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 & \cos\beta_1 \cos\alpha_3 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 & \cos\beta_2 \cos\alpha_3 \\ \cos\beta_3 \cos\alpha_1 & \cos\beta_3 \cos\alpha_2 & \cos\beta_3 \cos\alpha_3 \end{bmatrix} \begin{bmatrix} \cos\alpha_1 \\ \cos\alpha_2 \\ \cos\alpha_3 \end{bmatrix} \right)$$

$$= \left(\begin{bmatrix} \cos\beta_1 & 0 & 0 \\ 0 & \cos\beta_2 & 0 \\ 0 & 0 & \cos\beta_3 \end{bmatrix} \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 & \cos\beta_1 \cos\alpha_3 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 & \cos\beta_2 \cos\alpha_3 \\ \cos\beta_3 \cos\alpha_1 & \cos\beta_3 \cos\alpha_2 & \cos\beta_3 \cos\alpha_3 \end{bmatrix} \right) \begin{bmatrix} \cos\alpha_1 \\ \cos\alpha_2 \\ \cos\alpha_3 \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\beta_1 \cos\alpha_1 & \cos^2\beta_1 \cos\alpha_2 & \cos^2\beta_1 \cos\alpha_3 \\ \cos^2\beta_2 \cos\alpha_1 & \cos^2\beta_2 \cos\alpha_2 & \cos^2\beta_2 \cos\alpha_3 \\ \cos^2\beta_3 \cos\alpha_1 & \cos^2\beta_3 \cos\alpha_2 & \cos^2\beta_3 \cos\alpha_3 \end{bmatrix} \begin{bmatrix} \cos\alpha_1 \\ \cos\alpha_2 \\ \cos\alpha_3 \end{bmatrix}.$$

Из овога можемо извести супституцију

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \cos\beta_1 \\ \cos\beta_2 \\ \cos\beta_3 \end{bmatrix} = \begin{bmatrix} \cos\beta_1(\cos^2\alpha_1 + \cos^2\alpha_2 + \cos^2\alpha_3) \\ \cos\beta_2(\cos^2\alpha_1 + \cos^2\alpha_2 + \cos^2\alpha_3) \\ \cos\beta_3(\cos^2\alpha_1 + \cos^2\alpha_2 + \cos^2\alpha_3) \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 & \cos\beta_1 \cos\alpha_3 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 & \cos\beta_2 \cos\alpha_3 \\ \cos\beta_3 \cos\alpha_1 & \cos\beta_3 \cos\alpha_2 & \cos\beta_3 \cos\alpha_3 \end{bmatrix} \begin{bmatrix} \cos\alpha_1 \\ \cos\alpha_2 \\ \cos\alpha_3 \end{bmatrix},$$

односно $\vec{w} = \hat{Q}\vec{v}$, где је „матрица канала“ сада

$$\hat{Q} = \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 & \cos\beta_1 \cos\alpha_3 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 & \cos\beta_2 \cos\alpha_3 \\ \cos\beta_3 \cos\alpha_1 & \cos\beta_3 \cos\alpha_2 & \cos\beta_3 \cos\alpha_3 \end{bmatrix},$$

а улазни и излазни вектори, редом:

$$\vec{v} = \begin{bmatrix} \cos\alpha_1 \\ \cos\alpha_2 \\ \cos\alpha_3 \end{bmatrix} \quad \text{и} \quad \vec{w} = \begin{bmatrix} \cos\beta_1 \\ \cos\beta_2 \\ \cos\beta_3 \end{bmatrix}.$$

Коефицијенти $q_{ij} = \cos\beta_i \cos\alpha_j$ матрице $\hat{Q} = (q_{ij})$ нису условне вероватноће $P(\omega_i|\omega_j)$ у уобичајеном смислу, али приметимо да ни „условне вероватноће“ k_{ij} матрице канала $\hat{K} = (k_{ij})$ то нису биле обавезно. Ниједна од њих не мора бити класична „стохастичка матрица“.

Иначе, стохастичка матрица је квадратна матрица за описивање Ланаца Маркова чији коефицијенти могу представљати вероватноће. Такође се назива и матрица вероватноће, матрица прелаза, супституциона матрица, или Матрица Маркова. Има их три врсте:

1. Десна стохастичка матрица је права квадратна матрица, са сваким редом који се сабира на 1.
2. Лева стохастичка матрица је права квадратна матрица, са сваким ступцем укупног збира 1.
3. Двоструко стохастичка матрица је квадратна матрица ненегативних реалних бројева са сваким редом и ступцем који се збраја на 1.

Пример 1. Детерминанта матрице $\hat{Q}$ је нула, наиме:

$$\det \hat{Q} = \begin{vmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 & \cos\beta_1 \cos\alpha_3 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 & \cos\beta_2 \cos\alpha_3 \\ \cos\beta_3 \cos\alpha_1 & \cos\beta_3 \cos\alpha_2 & \cos\beta_3 \cos\alpha_3 \end{vmatrix} =$$

$$= \cos\alpha_1 \cos\alpha_2 \cos\alpha_3 \begin{vmatrix} \cos\beta_1 & \cos\beta_1 & \cos\beta_1 \\ \cos\beta_2 & \cos\beta_2 & \cos\beta_2 \\ \cos\beta_3 & \cos\beta_3 & \cos\beta_3 \end{vmatrix} = 0.$$

Заједнички фактор колоне ($\cos\alpha_j$) може се извући испред детерминанте. Затим, када су коефицијенти два ретка пропорционални ($\cos\beta_i$) детерминанта је нула. ❑

Матрица $\hat{Q}$, дакле, нема инверзну матрицу. Систем линеарних једначина $\vec{y} = \hat{Q}\vec{x}$ није регуларан, нема јединствено решење. Такви системи или немају решења уопште, или их имају бесконачно много (дуж неке праве, унутар равни и слично). У овом случају то је систем три од једне једначине и има:

- бесконачно много решења, ако је $\vec{y} = 0$;
- нема решења, ако је $\vec{y} \neq 0$.

Пример 2. Покажимо да систем:

$$\begin{cases} \cos\beta_1 \cos\alpha_1\, x_1 + \cos\beta_1 \cos\alpha_2\, x_2 + \cos\beta_1 \cos\alpha_3\, x_3 = c_1 \\ \cos\beta_2 \cos\alpha_1\, x_1 + \cos\beta_2 \cos\alpha_2\, x_2 + \cos\beta_3 \cos\alpha_3\, x_3 = c_2 \\ \cos\beta_3 \cos\alpha_1\, x_1 + \cos\beta_3 \cos\alpha_2\, x_2 + \cos\beta_3 \cos\alpha_3\, x_3 = c_3 \end{cases}$$

има бесконачно много решења ако $c_1 = c_2 = c_3 = 0$. Тада су пропорционалне сваке две (од три) датих једначина. Помножимо прву са $\cos\beta_2$ а другу са $\cos\beta_1$, па их одузмимо:

$$\cos\beta_2\, y_1 - \cos\beta_1\, y_2 = 0, \text{ а отуда } y_1 : y_2 = \cos\beta_1 : \cos\beta_2.$$

Слично можемо урадити са другом и трећом, или трећом и првом. Налазимо

$$y_1 : y_2 : y_3 = \cos\beta_1 : \cos\beta_2 : \cos\beta_3.$$

Ове једначине су заправо једна једначина писана на три начина. ❑

Трансформација $\hat{Q} : \vec{x} \to \vec{y}$ не памти оригинале и зато није типична „физичка“, јер не опонаша „закон одржања“. Пре него што пређемо на другачији модел „матрице канала“, погледајмо још неке особине ове.

03. Матрица канала $\hat{Q}$

Детерминанта матрице $\hat{Q}$ је нула [02. Пример 1] и, према томе, она нема инверзну матрицу ($\hat{Q}^{-1}$). Али, она зато има неке друге, једнако занимљиве, особине. Ради једноставности, посматрајмо је у једној равни, матрицу $\hat{Q}$ типа 2×2:

$$\hat{Q}\vec{v} = \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 \end{bmatrix} \begin{pmatrix} \cos\alpha_1 \\ \cos\alpha_2 \end{pmatrix} = \begin{pmatrix} \cos\beta_1 \\ \cos\beta_2 \end{pmatrix} = \vec{w}$$

јер је $\cos^2\alpha_1 + \cos^2\alpha_2 = 1$. Транспонујемо ли ову матрицу и множимо је другим вектором:

$$\hat{Q}^\tau\vec{w} = \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_2 \cos\alpha_1 \\ \cos\beta_1 \cos\alpha_2 & \cos\beta_2 \cos\alpha_2 \end{bmatrix} \begin{pmatrix} \cos\beta_1 \\ \cos\beta_2 \end{pmatrix} = \begin{pmatrix} \cos\alpha_1 \\ \cos\alpha_2 \end{pmatrix}$$

јер је $\cos^2\beta_1 + \cos^2\beta_2 = 1$. Дакле, транспонована матрица $\hat{Q}^\tau$ матрице $\hat{Q}$ понаша се као инверзна њој, иако она то није.

Пример 1. Налазимо производе матрице $\hat{Q}$ и њој транспоноване:

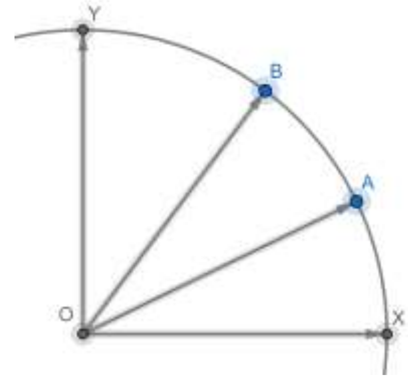
$$\hat{Q}\hat{Q}^\tau = \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 \end{bmatrix} \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_2 \cos\alpha_1 \\ \cos\beta_1 \cos\alpha_2 & \cos\beta_2 \cos\alpha_2 \end{bmatrix} =$$

$$= \begin{bmatrix} \cos^2\beta_1 & \cos\beta_1 \cos\beta_2 \\ \cos\beta_1 \cos\beta_2 & \cos^2\beta_2 \end{bmatrix} = \left(\hat{Q}^\tau\hat{Q}\right)^\tau.$$

Ово није јединична матрица за било које углове. ❑

На слици десно је Декартов правоугли систем *OXY* са јединичном кружницом центра у исходишту *O*. На кружници су две тачке *A* и *B* које дефинишу векторе $\vec{v} = \overrightarrow{OA}$ и $\vec{w} = \overrightarrow{OB}$. Вектори са (јединичним) векторима апсцисе $\vec{x}$ и ординате $\vec{y}$ заклапају углове: $\alpha_1 = \measuredangle(\vec{x}, \vec{a})$, $\beta_1 = \measuredangle(\vec{x}, \vec{b})$. Углови овима комплементни су: $\alpha_2 = 90° - \alpha_1$ и $\beta_2 = 90° - \beta_1$.

Дефиниције ових вектора помоћу пројекција на координатне осе су: $\vec{v} = (\cos\alpha_1, \cos\alpha_2)$ и $\vec{w} = (\cos\beta_1, \cos\beta_2)$. Прва и друга координата редом означавају апсцсу и ординату.



Из претходног знамо да матрица $\hat{Q}$ први трансформише у други ($\hat{Q}: \vec{v} \to \vec{w}$), а њој транспонована ради обрнуто, трансформише други вектор у први ($\hat{Q}^\tau: \vec{w} \to \vec{v}$). Међутим, из алгебре знамо да постоји и ротација око исходишта у равни *OXY*, за угао $\varphi = \beta_1 - \alpha_1$. Матрица ротације и њена транспонована су:

$$\hat{R}(\varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \quad \text{односно} \quad \hat{R}^\tau(\varphi) = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}$$

Прва, $\hat{R} = \hat{R}(\varphi)$, такође пресликава први вектор у други ($\hat{R}: \vec{v} \to \vec{u}$), а друга, њој транспонована матрица једнака је инверзној ($\hat{R}^\tau = \hat{R}^{-1}$). Заиста:

$$\hat{R}\hat{R}^\tau = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

што је производ који даје јединичну матрицу, једнако као производ $\hat{R}\hat{R}^{-1} = \hat{I}$. Транспоновањем ове добијамо и комутирану једнакост $\hat{R}^{-1}\hat{R} = \hat{I}$. О чему се овде ради појасниће следећи пример, надам се.

Пример 2. Разложимо матрицу ротације на два сабирка са основним угловима:

$$\hat{R} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} = \begin{bmatrix} \cos(\beta_1 - \alpha_1) & -\sin(\beta_1 - \alpha_1) \\ \sin(\beta_1 - \alpha_1) & \cos(\beta_1 - \alpha_1) \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\beta_1\cos\alpha_1 + \sin\beta_1\sin\alpha_1 & -\sin\beta_1\cos\alpha_1 + \cos\beta_1\sin\alpha_1 \\ \sin\beta_1\cos\alpha_1 - \cos\beta_1\sin\alpha_1 & \cos\beta_1\cos\alpha_1 + \sin\beta_1\sin\alpha_1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\beta_1\cos\alpha_1 + \cos\beta_2\cos\alpha_2 & -\cos\beta_2\cos\alpha_1 + \cos\beta_1\cos\alpha_2 \\ \cos\beta_2\cos\alpha_1 - \cos\beta_1\cos\alpha_2 & \cos\beta_2\cos\alpha_1 + \cos\beta_2\cos\alpha_2 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\beta_1\cos\alpha_1 & \cos\beta_1\cos\alpha_2 \\ \cos\beta_2\cos\alpha_1 & \cos\beta_2\cos\alpha_2 \end{bmatrix} + \begin{bmatrix} \cos\beta_2\cos\alpha_2 & -\cos\beta_2\cos\alpha_1 \\ -\cos\beta_1\cos\alpha_2 & \cos\beta_2\cos\alpha_2 \end{bmatrix}$$

$$= \hat{Q} + \hat{Q}_0.$$

Матрица $\hat{Q}_0$ није нулта, али делује ($\hat{Q}_0\vec{v} = \hat{0}$) као да јесте. ❑

Дакле имамо две необичне појаве, матрицу канала ($\hat{Q}{:}\vec{v} \to \vec{w}$) чија је детерминанта нула и, према томе, она није регуларна, нема инверзну матрицу, али њена транспонована ради инверзну трансформацију ($\hat{Q}^\tau{:}\,\vec{w} \to \vec{v}$). Друго, имамо не-нулту матрицу која ове векторе преводи у нула векторе ($\hat{Q}_0\vec{v} = \hat{0}$ и $\hat{Q}_0^\tau\vec{w} = \hat{0}$). Детерминанте обеју ових матрица су нуле.

Међутим:

$$\vec{w}^\tau\hat{Q}\vec{v} = (\cos\beta_1 \quad \cos\beta_2)\begin{bmatrix} \cos\beta_1\cos\alpha_1 & \cos\beta_1\cos\alpha_2 \\ \cos\beta_2\cos\alpha_1 & \cos\beta_2\cos\alpha_2 \end{bmatrix}\begin{pmatrix} \cos\alpha_1 \\ \cos\alpha_2 \end{pmatrix} =$$

$$= (\cos^2\beta_1\cos\alpha_1 + \cos^2\beta_2\cos\alpha_1 \quad \cos^2\beta_1\cos\alpha_2 + \cos^2\beta_2\cos\alpha_2)\begin{pmatrix} \cos\alpha_1 \\ \cos\alpha_2 \end{pmatrix}$$

$$= (\cos^2\beta_1\cos^2\alpha_1 + \cos^2\beta_2\cos^2\alpha_1) + (\cos^2\beta_1\cos^2\alpha_2 + \cos^2\beta_2\cos^2\alpha_2)$$

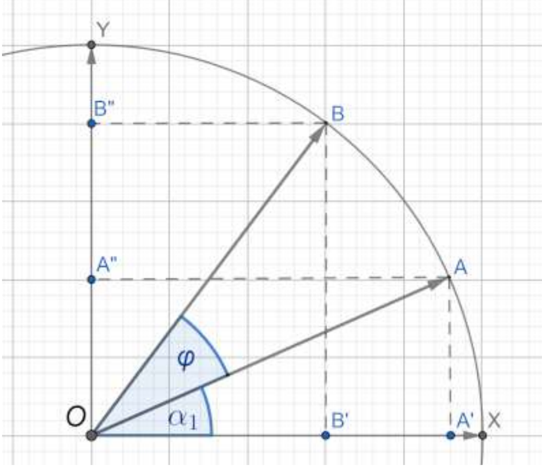
$$= (\cos^2\beta_1 + \cos^2\beta_2)\cos^2\alpha_1 + (\cos^2\beta_1 + \cos^2\beta_2)\cos^2\alpha_2$$

$$= \cos^2\alpha_1 + \cos^2\alpha_2 = 1.$$

Такође је: $\vec{v}^\tau\hat{Q}^\tau\vec{w} = \left(\vec{w}^\tau\hat{Q}\vec{v}\right)^\tau = 1$.

04. Површине

На слици је јединична кружница са два вектора, $\vec{a} = \overline{OA}$ и $\vec{b} = \overline{OB}$, угла између њих $\varphi = \sphericalangle(\vec{a}, \vec{b})$.



Пројекција тачке A на апсцису (x-осу) је A', а тачке B је B'. Угао $\sphericalangle(A'OA) = \alpha_1$, па због усклађивања са претходним текстом $\sphericalangle(A'OB) = \alpha_1 + \varphi = \beta_1$. Вектори $\vec{a}$ и $\vec{b}$ разапињу јединични ромб. Његова је површина производ дужина суседних страница и синуса угла захваћеног њима, дакле:

$$\mu(\vec{a}, \vec{b}) = |\vec{a}||\vec{b}| \sin\varphi =$$
$$= \sin\varphi = \sin(\beta_1 - \alpha_1)$$
$$= \sin\beta_1 \cos\alpha_1 - \cos\beta_1 \sin\alpha_1$$
$$= \cos\beta_2 \cos\alpha_1 - \cos\beta_1 \cos\alpha_2$$
$$= b_y a_x - b_x a_y = [\vec{a}, \vec{b}],$$

где $[\vec{a}, \vec{b}] = a_x b_y - b_x a_y = |\vec{a}||\vec{b}| \sin\varphi$ већ одавно називам комутатором вектора и, као што можете видети, он представља површину паралелограма којег дати вектори дефинишу. Називаћемо га и псеудо скаларним производом. Доследно претходном тексту, $\alpha_1 + \alpha_2 = \beta_1 + \beta_2 = 90°$.

Дужине вектора означавамо и са $a = |\vec{a}|$ и $b = |\vec{b}|$, када то не доводи до забуне, а пројекција:

$$a_x = \overline{OA}' = |\vec{a}| \cos\alpha_1 = a\cos\alpha_1 = \cos\alpha_1$$

$$a_y = \overline{OA}" = |\vec{a}| \cos\alpha_2 = a\cos\alpha_2 = \cos\alpha_2$$

и аналогно за $\vec{b}$. Обични и псеудо скаларни производи, сада знамо, су:

$$\begin{cases} \vec{a} \cdot \vec{b} = a_x b_x + a_y b_y = |\vec{a}||\vec{b}| \cos\varphi = ab\cos\varphi = \cos\varphi \\ |\overrightarrow{a} \times \overrightarrow{b}| = a_x b_y - a_y b_x = |\vec{a}||\vec{b}| \sin\varphi = ab\sin\varphi = \sin\varphi \end{cases}$$

Пример 1. Изведимо матрицу ротације $\hat{R} = \hat{R}(\varphi)$, коришћену у претходном наслову:

$$b_x = (a_x^2 + a_y^2)b_x + a_x a_y b_y - a_x a_y b_y =$$

$$= (a_x b_x + a_y b_y)a_x - (a_x b_y - a_y b_x)a_y$$

$$= (\cos\varphi)a_x - (\sin\varphi)a_y$$

и аналогно $b_y = (\sin\varphi)a_x + (\cos\varphi)a_y$. Матрично добисмо

$$\begin{pmatrix} b_x \\ b_y \end{pmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{pmatrix} a_x \\ a_y \end{pmatrix}$$

или краће $\vec{b} = \hat{R}\vec{a}$, а то је тражена формула ротације. ❑

То је занимљива употреба ове „методе површина", као и следећа. Обични и псеудо скаларни производ вектора генералишимо помоћу функције „косинус и синус" $\text{cis}\, x = \cos x + i\sin x$. Приметимо, $\text{cis}\, x = e^{ix} = \exp ix$, где за имагинарну јединицу важи $i^2 = -1$. Тада су корисне

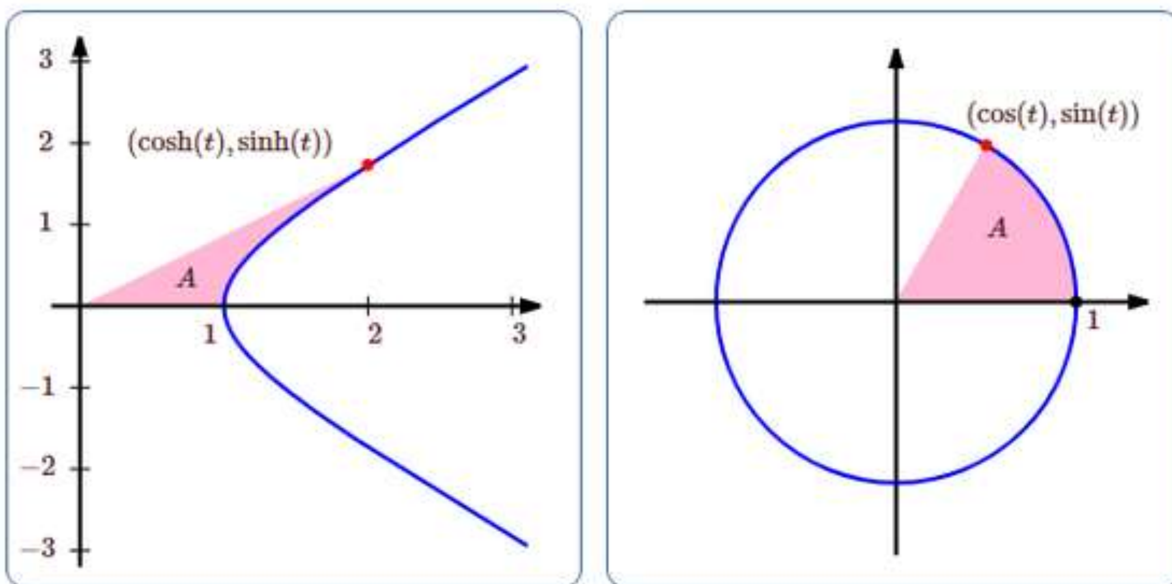
тригонометријске „хиперболне функције“ $\cosh x = (e^x + e^{-x})/2$ и $\sinh x = (e^x - e^{-x})/2$, посебно због веза са обичним:

$$\cos ix = \cosh x \quad \text{и} \quad \sin ix = i \sinh x.$$

Отуда, на пример $\text{cis}\, i\varphi = \cosh\varphi - i\sinh\varphi$, и даље доследна примена у аналитичкој геометрији троугла, или посебно овде у матрицама ротације

$$\hat{R}(i\varphi) = \begin{bmatrix} \cos i\varphi & -\sin i\varphi \\ \sin i\varphi & \cos i\varphi \end{bmatrix} = \begin{bmatrix} \cosh\varphi & -i\sinh\varphi \\ i\sinh\varphi & \cosh\varphi \end{bmatrix}.$$

У хиперболној геометрији постоји и одговарајућа „јединична кружница“. Двоструко од шрафиране површине захваћене углом t на хиперболи лево одговарајућа је на кружници десно (Hiperbolic).



Приметимо да у овој, иначе познатој интерпретацији, угао $t = 2A$ у радијанима представља двоструку површину (Area) кружног исечка десно (шрафирано) и да она постаје одговарајућа променљива хиперболних функција лево. Наглашавам, двоструким повећањем шрафиране површине лево неће се удвостручити „угао“ у смислу углова на слици десно, већ остају само аналогије ове две геометрије. Детаљније о овоме можете наћи и у мојој скрипти Тригонометрија – равна, сферна и хиперболна, писаној 2016. године за боље ђаке Гимназије Бања Лука.

Навикли смо да у класичној тригонометрији угао у радијанима сматрамо „дужином“ лука јединичне кружнице који се из центра види под датим углом, али се та идеја не преноси на хиперболу. Двема тригонометријама заједничко је тумачење угла површином и ту сада наступа „теорија информације“.

05. Хартлијева информација

Теорија физичке информације треба да има упориште у ономе што сматрамо нашим опажајима (Senses), затим у закону одржања (Conservation), а коначно и у математичким апстракцијама. Овим трећима још увек је добро полазиште Хартлијева (1928) дефиниција информације (Information).

Бинарно претраживање $n = 1, 2, 3, ...$ равноправних елемената своди се на постављање питања која почињу са „Да ли ...?“ а на које испитивани одговара једним од „Да“ или „Не“. На пример, замишљен је један од $n = 8$ бројева, рецимо број 7 у почетку непознат испитивачу. Прво од $h = 3$ његових питања било би. „Да ли се тражени број налази у првој половини свих могућих?“. Како он није у скупу бројева {1, 2, 3, 4}, одговор ће бити „Не“.

Следи друго питање: „Да ли се тражени број налази у првој половини преосталих могућих?“. Он није ни у том скупу, јер 7 ∉ {5, 6}, па је одговор опет „Не“. Треће и у овом случају последње питање било би „Да ли је тражени број 7?“ и одговор је „Да“. Била су потребна три (бинарна) питања, за $2^3 = 8$ могућности, па кажемо (Хартлијева) инфомација 8 једнако вероватних исхода износи $\log_2 8 = 3$. Уопште, за $n = 2^h$ равноправних исхода та је информација $h = \log_2 n$.

Дубину течне хране опажамо на основу логаритамских стимулуса (Logarithmic Perception), на сличан начин на који перципирамо визуелне или слушне информације (Weber–Fechner law). Укратко, можемо сматрати да Хартлијева информација довољно приближно одговара нашим опажајима. Поред тога, потреба комуницирања путем граница, са ове и оне стране субјекта односно објекта, говори о важности површина у тим процесима и посебно о вези преноса информације и физичког дејства. Коначно, закон одржања информације више пута сам расправљао и налазио га оправданим.

Бинарно писаних бројева са пет цифара је 32 (од 0 до 31). У систему бројева базе три ($b = 3$) до највише петоцифрених бројева је 243 $243 = 3^5$, базе 10 је 10 000, а уопште $n = b^h$ је h-цифрених бројева система базе b. Важи и обрнуто, скуп n једнако изгледних опција има Хартлијеву информацију $h = \log_b n$ у јединицама информације дефинисаним базом b логаритма. Са $b = 2$ та јединица је „бит“, са $b = 10$ јединица информације је „децит“, а за ($e \approx 2{,}71828$) базу природног логаритма та је јединица „нат“ (скраћено: *natural*).

Са овим базама можемо ићи до крајњих граница природних бројева, када $h \to \aleph_0$, тада $n \to b^{\aleph_0}$, за произвољну базу $b \in \{2,3,4, ...\}$ којим се иначе може описивати исти скуп природних бројева $\mathbb{N}$. Алеф нула је ознака за „пребројиву бесконачност“ природних, целих и рационалних бројева, а $2^{\aleph_0}$ за већу бесконачност, рецимо ирационалних и реалних бројева, коју називамо „континуум“.

Са генерализацијом Хартлијеве информације иде се и даље, са бројем могућности било којим позитивним реалним бројем ($n \in \mathbb{R}$), једноставно зато што то дозвољава њена форма $h = \log n$. У мојој теорији информације чини се и корак даље у генерализацији Хартлијеве формуле на комплексне бројеве ($n \in \mathbb{C}$), а и то једноставно дозвољава иста форма $h = \log n$, као и инверзна $n = e^h$.

Приметимо да стављајући $h = i\varphi$ и прихватајући идеју „угла“ као површине, саму информацију разумемо као површину, а због периодичности логаритма комплексног броја, као периодичну

појаву. Посебно, када такав угао схватимо као „дејство“ (производ импулса и пута, односно енергије и времена), ето нас у квантно механичком тумачењу таласне функције као информације (Schrödinger equation). Тада је ова Хартлијева информација еквивалентна рецимо $kx - \omega t$.

Обзиром да се у експоненту ови поједини фактори $-px/\hbar$, односно $-Et/\hbar$, могу сабирати и формирати изразе попут скаларног производа $(xp_x + yp_y + zp_z - tE)$, генерализације Хартлијеве информације довести ће нас и до „информације перцепције“. Тако помоћу теорије информације добијамо нови увид у квантне процесе, а за узврат (Ема Нетер) добијамо један од доказа закона одржања Хартлијеве, наизглед чисто математичке форме информације. Такође, анализирајући разматрање Ландауера (1961) на начин у „квантниом рачуну“ приметићемо да нема креације, емисије, преноса, или копирања информације без потрошње, односно размене енергије.

Да нису само „квантна стања“, или честице-таласи, дакле репрезентације вектора Хилбертове алгебре, неке информације, већ су то и „квантни процеси“ који су репрезентације линеарних оператора, видимо прво из саме дефиниције вектора, односно векторских простора (Квантна механика). Међутим, исто разумемо и интуитивно. Електрон у интеракцији са околином пролази кроз себи својствене процесе, који га дефинишу и по чему знамо да је управо реч о електрону, чиме заправо откривамо дуализам стања и процеса иначе детаљније познат у односу вектора и одговарајућих оператора који на те векторе делују.

Тиме долазимо до даље природне генерализације Хартлијеве информације. Сада до саме његове поопштене „информације перцепције“ $S = A_1B_1 + A_2B_2 + A_3B_3 + \cdots$ у којој ће A_k и B_k (редом за индексе $k = 1, 2, 3, \ldots$) бити линеарни оператори, а не пуки вектори. Посебно, квадратне форме су врсте информације перцепције.

Посебна прича те генерализације је адаптација. То је спонтани, општи али благи процес „бирања лакшег пута“ (метафорички, као да природа има „свест“ и нешто „бира“), који у екстремном случају препознајемо на мртвој твари, у облику физичког принципа најмањег дејства, а који тек требамо разумети на системима веће акомулиране информације, попут живих бића, који зато имају веће изборе.

Јединке се удружују у колективе предајући своје слободе (количине опција) друштву, пратећи „принцип мање информације“, односно „принцип веће вероватноће“ реализујући чешће вероватнија стања. Оне то покушавајући несвесно учествују такође у грађењу сложенијих живих бића на основу једноставнијих, поред покушаја да слободу предају и непосредно мртвој твари. Наиме, Хартлијева информација је мања када је вероватноћа већа, следи из $h = \log n = -\log\frac{1}{n} = -\log P(n)$, јер је вероватноћа појединог $P(n)$ од n једнако вероватних исхода то мања што је тих исхода више.

Међутим, процеси су такође стања (вектори и оператори који их трансформишу дуалне су појаве), па ћемо исту адаптацију препознавати и самим процесима, рецимо у деформацији сигнала проласком кроз низ канала (Марковљеви ланци) која углавном конвергира ка сопственој вредности мултиплициране матрице канала.

06. Сопствене величине матрице $\hat{Q}$

Својствене вредности (λ) и својствене векторе (x), за које важи матрична једначина $\hat{M}x = \lambda x$, кратко називамо карактеристичним, или сопственим величинама.

Пример 1. У случају опште матрице 2×2 налазимо:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

$$\det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} = 0$$

$$(a-\lambda)(c-\lambda) - bc = 0$$

$$\lambda^2 - (a+c)\lambda + ad - bc = 0.$$

Својствене вредности (eigenvalue) су решења ове квадратне једначине по λ, а познајући их откривамо и припадне својствене векторе (eigenvector). ❑

Идеја је да се нађе стање (својствени вектор x) који процес (матрица $\hat{M}$) не мења, осим што га помножи својственом вредношћу λ. То је познато, али оно из следећа два примера није.

Пример 2. Нађимо сопствене вредности матрице $\hat{Q}$:

$$0 = \det(\hat{Q} - \lambda \hat{I}) = \begin{vmatrix} \cos\beta_1 \cos\alpha_1 - \lambda & \cos\beta_1 \cos\alpha_2 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 - \lambda \end{vmatrix} =$$

$$= (\cos\beta_1 \cos\alpha_1 - \lambda)(\cos\beta_2 \cos\alpha_2 - \lambda) - \cos\beta_1 \cos\beta_2 \cos\alpha_1 \cos\alpha_2$$

$$= \lambda^2 - (\cos\beta_1 \cos\alpha_1 + \cos\beta_2 \cos\alpha_2)\lambda$$

$$= \lambda^2 - (\cos\beta_1 \cos\alpha_1 + \sin\beta_1 \sin\alpha_1)\lambda$$

$$= \lambda^2 - \cos(\beta_1 - \alpha_1)\lambda$$

$$= \lambda(\lambda - \cos\varphi),$$

а отуда главна својствена вредност $\lambda_1 = \cos\varphi$, док друго решење $\lambda_2 = 0$ одбацујемо. ❑

Пример 3. Квадрирањем матрице $\hat{Q}$ добијамо:

$$\hat{Q}^2 = \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 \end{bmatrix}^2 =$$

$$= \begin{bmatrix} \cos^2\beta_1 \cos^2\alpha_1 + \cos\beta_1 \cos\beta_2 \cos\alpha_1 \cos\alpha_2 & \cos^2\beta_1 \cos\alpha_1 \cos\alpha_2 + \cos\beta_1 \cos\beta_2 \cos^2\alpha_2 \\ \cos\beta_1 \cos\beta_2 \cos^2\alpha_1 + \cos^2\beta_2 \cos\alpha_1 \cos\alpha_2 & \cos\beta_1 \cos\beta_2 \cos\alpha_1 \cos\alpha_2 + \cos^2\beta_2 \cos^2\alpha_2 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\beta_1 \cos\alpha_1 (\cos\beta_1 \cos\alpha_1 + \cos\beta_2 \cos\alpha_2) & \cos\beta_1 \cos\alpha_2 (\cos\beta_1 \cos\alpha_1 + \cos\beta_2 \cos\alpha_2) \\ \cos\beta_2 \cos\alpha_1 (\cos\beta_1 \cos\alpha_1 + \cos\beta_2 \cos\alpha_2) & \cos\beta_2 \cos\alpha_2 (\cos\beta_1 \cos\alpha_1 + \cos\beta_2 \cos\alpha_2) \end{bmatrix}$$

$$= (\cos\beta_1 \cos\alpha_1 + \cos\beta_2 \cos\alpha_2)\begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 \end{bmatrix} = \cos(\beta_1 - \alpha_1)\ \hat{Q}.$$

Дакле, $\hat{Q}^2 = \cos\varphi\ \hat{Q}$, а затим за $n = 1, 2, 3, \dots$ добијамо $\hat{Q}^n = \cos^{n-1}\varphi\,\hat{Q}$. ❑

На основу ових примера и дефиниције сопствених величина ($\hat{Q}x = \lambda x$ и $\lambda = \cos\varphi$) налазимо:

$$\hat{Q}^n x = \hat{Q}^{n-1}(\hat{Q}x) = \hat{Q}^{n-1}(\lambda x) = \lambda(\hat{Q}^{n-1}x) = \cdots = \lambda^n x,$$

$$\hat{Q}^n x = \cos^n\varphi\, x.$$

Својствена вредност n-тог степена матрице $\hat{Q}$ је $\cos^n\varphi$. Када је $0 < |\cos\varphi| < 1$, а углавном је, тада степеновањем ова својствена вредност тежи нули, $\lim_{n\to\infty}\cos^n\varphi = 0$.

Када би порука својственим вектором пролазила композицијом канала матрице $\hat{Q}$ излазила би све слабија. Таква би била обзервабла, односно шанса опажања дате мерљиве физичке величине, коју сопствена вредност представља. Погледајмо сада шта је својствени вектор који одговара тој својственој вредности.

Пример 4. Решавањем матричне једначине $(\hat{Q} - \hat{I}\cos\varphi)x = 0$ налазимо и проверавамо:

$$\begin{cases} (\cos\beta_1\cos\alpha_1 - \cos\varphi)x_1 + \cos\beta_1\cos\alpha_2\, x_2 = 0 \\ \cos\beta_2\cos\alpha_1\, x_1 + (\cos\beta_2\cos\alpha_2 - \cos\varphi)x_2 = 0 \end{cases}$$

$$\begin{cases} (\cos\beta_1\cos\alpha_1 - \cos(\beta_1 - \alpha_1))x_1 + \cos\beta_1\cos\alpha_2\, x_2 = 0 \\ \cos\beta_2\cos\alpha_1\, x_1 + (\cos\beta_2\cos\alpha_2 - \cos(\beta_1 - \alpha_1))x_2 = 0 \end{cases}$$

$$\begin{cases} (\cos\beta_1\cos\alpha_1 - \cos\beta_1\cos\alpha_1 - \sin\beta_1\sin\alpha_1)x_1 + \cos\beta_1\cos\alpha_2\, x_2 = 0 \\ \cos\beta_2\cos\alpha_1\, x_1 + (\cos\beta_2\cos\alpha_2 - \cos\beta_1\cos\alpha_1 - \sin\beta_1\sin\alpha_1)x_2 = 0 \end{cases}$$

$$\begin{cases} -\sin\beta_1\sin\alpha_1\, x_1 + \cos\beta_1\cos\alpha_2\, x_2 = 0 \\ \ \cos\beta_2\cos\alpha_1\, x_1 - \cos\beta_1\cos\alpha_1\, x_2 = 0 \end{cases}$$

$$\begin{cases} -\cos\beta_2\, x_1 + \cos\beta_1\, x_2 = 0 \\ \ \cos\beta_2\, x_1 - \cos\beta_1\, x_2 = 0 \end{cases}$$

$$\cos\beta_2\, x_1 = \cos\beta_1\, x_2$$

$$x_1 = \cos\beta_1,\ \ x_2 = \cos\beta_2$$

$$x = \begin{pmatrix} \cos\beta_1 \\ \cos\beta_2 \end{pmatrix}.$$

То је својствени вектор матрице $\hat{Q}$ који одговара својственој вредности $\lambda = \cos\varphi$. ❑

07. Матрица канала $\hat{K}$

Теорија информације рачуна са спрегом субјекта и објекта. Тако је канал $\hat{Q} = \hat{Q}(\alpha, \beta)$ осмишљен како би јединични вектор $u(\alpha)$ нагиба α према првој оси променио тај свој угао у β и постао јединични вектор $u(\beta)$. Процеси зависе од стања на која делују и обрнуто. То је природно.

Стања, дакле врсте квантних честица-таласа препознајемо у експериментима управо према траговима (чит. процесима) које следе. Стања су пробирљива према процесима и обрнуто; неће се свако стање подвргнути сваком процесу. Задржимо се мало на овом запажању на примеру матрице $\hat{Q}$.

Пример 1. Применимо матрицу $\hat{Q}(\alpha, \beta)$ на јединични вектор $u(\gamma)$:

$$\hat{Q}(\alpha, \beta)\ u(\gamma) = \begin{bmatrix} \cos\beta_1 \cos\alpha_1 & \cos\beta_1 \cos\alpha_2 \\ \cos\beta_2 \cos\alpha_1 & \cos\beta_2 \cos\alpha_2 \end{bmatrix} \begin{pmatrix} \cos\gamma_1 \\ \cos\gamma_2 \end{pmatrix} =$$

$$= \begin{bmatrix} \cos\beta \cos\alpha & \cos\beta \sin\alpha \\ \sin\beta \cos\alpha & \sin\beta \sin\alpha \end{bmatrix} \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix}$$

$$= \begin{bmatrix} \cos\beta \cos\alpha \cos\gamma + \cos\beta \sin\alpha \sin\gamma \\ \sin\beta \cos\alpha \cos\gamma + \sin\beta \sin\alpha \sin\gamma \end{bmatrix}$$

$$= (\cos\alpha \cos\gamma + \sin\alpha \sin\gamma) \begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix}$$

$$= \cos(\alpha - \gamma)\ u(\beta).$$

Резултат је вектор $u(\beta)$ умањен $\cos(\alpha - \gamma)$ пута. ❑

Из примера видимо да матрица $\hat{Q}$ није матрица канала $\hat{K}$ у класичном смислу (која би сваки вектор расподеле вероватноћа преводила у неки вектор расподеле), следећу у расправи. За сада приметимо само да код ове ($\hat{Q}$) није неопходно нормирање резултата, које би резултату пресликавања њоме давало вектору $u(\beta)$ увек исти нагиб β, обзира на улазни угао γ почетном. То би, пак, био уобичајени поступак са векторима квантне механике (ψ).

Потреба нормирања резултата указује нам на немогућност примене датог процеса на произвољна стања, а одсуство комплексних бројева на игнорисање паралелних реалности. Поменуте две врсте „матрица канала“, које су реалне и не морају бити квадратне, треба разликовати од унитарних, $\hat{U}$, квантне механике. У овом наслову бавимо се даље класичном матрицом канала $\hat{K} = \left(k_{ij}\right)$, чији су коефицијенти условне вероватноће, а улазни и излазни вектори расподеле вероватноћа.

Нека је $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ потпуни скуп неких $n = 1, 2, 3, \dots$ случајних догађаја који чине једну расподелу вероватноћа. Тада је вероватноћа производа, пресека различитих $P\left(\omega_i \cap \omega_j\right) = 0$, а збир свих $\sum_k P(\omega_k) = 1$. Вероватноћа слања j-тог сигнала је $p_j = P\left(\omega_j\right)$, а q_i вероватноћа пријема i-тог сигнала на излазу, након проласка каналом матрице $\hat{K}$. Вероватноћа да је послат j-ти а примљен i-ти сигнал је k_{ij}. Збирове (уније), или производе (пресеке) ових елементарних догађаја (ω_k) означаваћемо великим словима A, B, C и сличним.

Формула потпуне вероватноће каже, као што је познато, да је $q_i = \sum_{j=1}^{n} k_{ij}\ p_j$, или матрично

$$q = \hat{K}p = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$$

у случају $n = 3$ (Channel).

Иначе је условна вероватноћа $P(A|B) = P(AB)/P(B)$, тј. вероватноћа да ће се десити исход A ако се десио B. Увек је $P(A|B) \geq 0$, дакле вероватноћа је ненегативан реалан број за произвољне догађаје $A, B \subseteq \Omega$. Вероватноћа бар некога од свих, ако се десило било које B, сигуран је догађај, $P(\Omega|B) = 1$. То су последице прве и друге аксиоме теорије вероватноће. Из треће и последње аксиоме, да је вероватноћа збира (уније) дисјунктних догађаја A_j (једног или више ω_i) једнака збиру њихових вероватноћа, следи за условне вероватноће $P\left[\left(\sum_j A_j\right)|B\right] = \sum_j P\left(A_j|B\right)$.

Теорема 1. Када је $p = [\![p_j]\!]$ вектор произвољне расподеле вероватноћа (дисјунктних исхода и извесних да ће се један од њих десити), онда је $q = \hat{K}p$ такође вектор неке расподеле.

Доказ: Како су сви коефицијенти $k_{ij}\, p_j$ ненегативни то је и њихова линеарна комбинација ненегативна. Даље због $q_i = \sum_{j=1}^{n} k_{ij}\, p_j$ имамо:

$$\sum_{i=1}^{m} q_i = \sum_{i=1}^{m} \sum_{j=1}^{n} k_{ij} p_j = \sum_{j=1}^{n} p_j \left(\sum_{i=1}^{m} k_{ij}\right) = \sum_{j=1}^{n} p_j = 1,$$

тј. $\sum_{i=1}^{m} q_i = \sum_{j=1}^{n} p_j = 1$, што их чини расподелама. ■

У доказу ове теореме приметимо да је матрица канала $\hat{K}$ типа $m \times n$ и да канал не мора бити квадратни. Пример таквих је „Семафор“. Ова теорема се користи и за доказивање егзистенције канала.

Ово је тзв. хоризонтална матрица канала, јер њен редак (врсту) множимо са колоном расподеле, која је контраваријанти вектор. Ако је вектор p расподеле (n-компонентни) коваријантан, представљен врстом (хоризонтално), онда је одговарајућа матрица канала „вертикална“ (тада типа $n \times m$), а резултирајући вектор q (m-члан низ) хоризонталан.

Теорема 2. Матрични производ произвољних хоризонталних матрица канала, ако постоји, је хоризонтална матрица канала.

Доказ: Да би матрични производ био могућ, прва матрица мора имати онолико колона (r) колико друга има врста. Како је сваки коефицијент датих матрица ненегативан, таква је и њихова линеарна комбинација, $c_{ij} = \sum_{k=1}^{r} a_{ik}\, b_{kj}$, па за свако $i = 1, 2, \dots, m$ је:

$$\sum_{j=1}^{n} c_{ij} = \sum_{j=1}^{n} \sum_{k=1}^{r} a_{ik} b_{kj} = \sum_{k=1}^{r} a_{ik} \sum_{j=1}^{n} b_{kj} = \sum_{k=1}^{r} a_{ik} = 1.$$

Дакле, сваки i-ти редак производа матрица ($\hat{C}_{m\times n} = \hat{A}_{m\times r}\hat{B}_{r\times n}$) чини неку расподелу. ■

08. Адаптација каналом

Канал какав год да је ($\hat{K}$, $\hat{Q}$, или $\hat{U}$) може мењати улазне податке, процесуирати их, кварити или адаптирати и претварати у излазне. Неконтролисано ометање, расипање и губитак података је сметња или „шум“, а конвергенција ка неком одређеном типу излаза је „адаптација“. У прошлом наслову [07. Пример 1] видели смо један гранични пример. То слабљење свих излазних сигнала између је нетипичне адаптације и правог шума.

Други пример покушајмо разумети на начин како је објашњаван „начелни минимализам“ информације (Minimalism). Погледајмо то исто помоћу „информације перцепције“, рецимо облика $S = a_1b_1 + \cdots + a_nb_n$ где су низови расподеле вероватноћа. Нека је:

$$m_a = \min_{1\le k\le n} a_k,\ m_b = \min_{1\le k\le n} b_k,\ M_a = \max_{1\le k\le n} a_k,\ M_b = \max_{1\le k\le n} b_k$$

$$m = \max\{m_a, m_b\},\ M = \min\{M_a, M_b\},$$

па приметимо да је $m \le S \le M$. У поменутом прилогу, о минимализму, ово је доказано и кориштено за запажање да се распон вероватноћа расподеле смањује након проласка каналом, осим у случају када сваки ред матрице ($\hat{K}$) има јединицу и, према томе, све остале нуле. Другим речима, иако је регуларна (детерминанта јој није нула), матрица канала имаће инверзну само ако су њене врсте (колоне) пермутације јединичне матрице.

Даље, овде примећујемо, матрица канала која није настала пермутацијама јединичне, смањује опсег (дисперзију) вероватноћа улазне расподеле, која ће зато вишеструким понављањем проласка кроз исти канал конвергирати. Није тешко препознати Кошијев низ и конвергенцију у том случају чија је гранична вредност својствени вектор. То је „адаптација каналом“.

Адаптација каналом је позната појава из Марковљевих ланаца. Она је тамо на другачији начин доказивана, а интерпретирана је доследно класичној теорији информације и комуникације. Међутим, сличне је форме и заједничких дубљих узрока са онима из ове „теорије информације“, као спонтана тежња стохастичких система ка развоју у мање информативне.

Слично се догађа са субјектима и објектима опажања у оквиру саме информације перцепције. Они се једни другима прилагођавају тако да је тај скаларни производ вектора ($S = a \cdot b$) што мањи, обзиром на задате њихове интензитете, што значи адаптацију фактора у сабирцима тако да мањи иде са већим и обрнуто, већи са мањим. Донекле се слично „прилагођавање“ догађа и са копијом вектора у односу на ретке (хоризонталне) матрице канала.

Из опште матрице другог реда [06. Пример 1] налазимо средњу својствену вредност $\lambda = (a + c)/2$, а онда за компоненте својственог вектора $x_1 : x_2 = b : (a - \lambda) = (\lambda - d) : c$. Видимо да координате својственог вектора теже расти са супротним коефицијентима колона. Прва (x_1) са другим (b), а друга (x_2) са првим (c). Ове тенденције компликују се са осталим, због чега је ову обрнуту размеру својствених вектора било тако тешко приметити и ненаглашавати у литератури.

Непримећена је, или је барем ненаглашена, остајала и блага али општа тежња физичких система да еволуирају ка мање информативнима. Њу коју компликује свеприсутност информације и њен закон одржања, такође тек од недавно (ако и тада) примећени у „теорији информације“.

Сличност између субјеката учесника адаптације информације перцепције (S) и адаптације вектора расподеле каналима преноса ($\widehat{K}$) је узајамно прилагођавање субјеката и објеката комуникације. Ми бирајући опције мењамо и будуће своје боравиште, односно бирамо процесе којима можемо располагати. Разлика ових адаптација је у већој објективности канала преноса наспрам простих саговорника, у немогућности да саме канале тако лако мењамо колико да их одбацујемо.

На описани начин се може приметити једног другачијег начела информације, од поменутог минимализма. То је потреба за разноврсношћу, тежња природе ка различитостима. Када би се сам канал преноса вишеструком применом мењао тако да му се распон вероватноћа у ретцима (колонама) смањује, то би убрзавало процес конвергенције полазних сигнала ка својственом вектору. Убрзано би се смањивале дисперзије овог света.

Са друге стране, као супротност овом гушењу различитости стоје матрице канала чије су врсте (колоне) пермутације јединичне матрице и чије су детерминанте јединичне. Детерминанте дисперзивних канала, које у врстама немају по једну јединицу и све остале нуле, такође су различите од нуле, али су мање од јединице и зато описују једносмерне процесе (Minimalism). Наиме, сужавањем распона вероватноћа расподеле након њеног проласка каналом, њено враћање назад инверзном матрицом (која постоји, јер јој је детерминанта различита од нуле) неће бити могуће, неће довести до истог оригинала.

Зато су унитарне матрице (чије детерминанте су јединичне) толико важне, пре свега у свету микро физике, где су важне и вероватноће. У макро свету доминирају закони великих бројева теорије вероватноће, у квантној механици не толико, па су повећане и шансе пребега стања из „стварности већине“ у за њих неку „псеудо реалност“. Зато су унитарне матрице квантне физике тако препуне комплексних коефицијената, чије присуство указује на могућности „заобилажења“.

Укратко речено, поред принципа најмањег дејства, дакле минимализам информације, природом владају и потребе за разликама. Први је последица, тачније еквивалент „максимализма вероватноће“ (већој вероватноћи одговара мања информација), а други је последица „несталности информације“ (поновљена „вест“ више није вест). Ове две аспирације противне су једна другој и уравнотежују се. Из ове друге настаје и потреба за већом „објективношћу“ (чит. непромењивошћу) неких учесника комуникације, а затим и отпор уједначавању свега и свачега, те различитости.

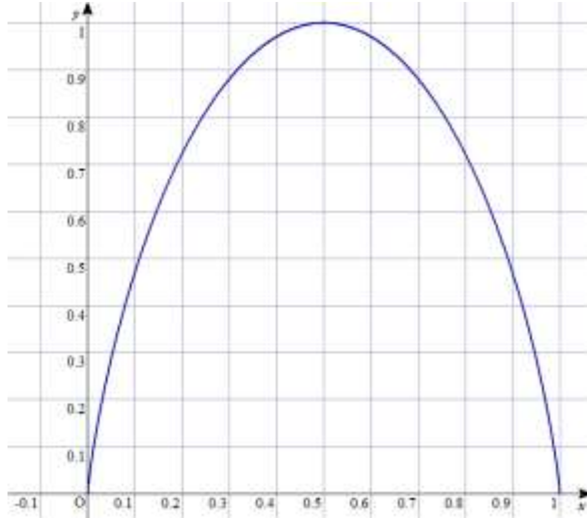
09. Различитост и закон одржања

Видели смо [08. Адаптација] како потреба за различитошћу и принцип минимализма контрирају једно другоме, а сада ћемо погледати везе те принудне различитости и закона одржања. Увек се ради о информатичкој основи.

Пре свега, приметимо да промена дисперзије расподеле вероватноће мења информацију. На пример, у мојој књизи „Физичка информација“ наћи ћете граф функције

$$y(x) = -x\log_2 x - (1-x)\log_2(1-x)$$

као на слици лево. То је Шенонова бинарна информација, где је $x \in (0,1)$ вероватноћа да ће се неки случајан догађај десити, а $1-x$ је вероватноћа да се исти неће десити. Ордината $y = y(x)$ представља информацију у битима.



Као што се види, највећу информацију, $y = 1$, има (бинарни) догађај вероватноће $x = 1/2$. То је ситуација једнаких вероватноћа исхода, да се дати деси и да се не деси.

У сложенијем случају, $n \in \mathbb{N}$, могућих исхода вероватноћа $p_1, p_2, \dots, p_n$, које су позитивни бројеви ($p_k > 0$, за свако $k = 1, 2, \dots, n$) јединичног збира ($p_1 + p_2 + \dots + p_n = 1$), Шенонова информација је $S(p_1, p_2, \dots, p_n) = -p_1 \log_b p_1 - p_2 \log_b p_2 - \dots - p_n \log_b p_n$, где је база логаритма произвољан број $b > 0$ који дефинише јединицу мере информације.

У једном питању на блогу (Emergence II: Можете ли ми још једном објаснити настајање својстава ентитета којих његови делови сами по себи немају?), доказивао сам да производ низова, скалани производ вектора, као је и Шенонова информација, мења вредност зависно од узајамног поретка. Када се сабирци, а тиме и фактори у њима, могу поредати тако да су оба низа исте монотоности (оба растућа, или оба опадајућа), онда је њихов збир максималан, а у обрнутом случају, када је један од низова растући а други опадајући, онда је збир множених парова минималан.

Са већом вероватноћом p, мања је информација $-\log p$, па је Шенонова информација најмања те врсте. Вероватноће дате расподеле (p_k) множене логаритмима вероватноћа друге расподеле (q_k) не би могле бити мање од Шенонове ($-\sum p_k \log p_k \le -\sum p_k \log q_k$). Посебно, ставимо ли да је та друга расподела униформна ($q_k = 1/n$, за свако $k = 1, 2, \dots, n$) важиће исто. Дакле, Шенонова информација постиже максимум када су сви исходи једнако вероватни.

У поменутој (мојој) књизи, разрађиване су још неке врсте дефиниција информације. Оне су блиске Шеноновој, али су дефинисане тако да за њих важи закон одржања и такође расту уједначавањем вероватноћа исхода. Из тога извлачимо поуку да природа не може дозволити превише једнакости да не би нарушила закон одржања. Штавише, у једном одговору у блогу (Channel, Теорема) наћи ћете доказ да је (апсолутна) вредност сваког дијагоналног елемента произвољне матрице већа (или да је свака мања) од збира осталих довољан услов да детерминанта те матрице није нула.

Примењено на матрице канала, тај став значи да је различитост довољан услов закона одржања информације. Са принципом минимализма, који лако произилази из начелно чешћег догађања

вероватнијих исхода, затим мање информације вероватнијих, те објективности случајности, са много већом убедљивошћу налазимо да природа не воли једнакост. Тај став зато не изненађује нас у теорији информације.

Али он је прилично чудно откриће за филозофију и савремену науку. Потврде ипак налазимо у многострукостима света који нас окружује, а јаче његово упориште могли бисмо препознати у Паулијевом принципу искључења (1925). Према том „никоја два идентична фермиона не могу да се налазе у истом квантном стању симултано“. Фермиони су честице супстанце, за разлику од бозона који дозвољавају слагање истих на истом месту (попут простора). Међутим, генерализација Паулијевог искључења на ширу супстанцу чинила се предалеким циљем.

Посебно би нам могла бити изненађујућа дедукција, да из различистости следи закон одржања. Ако је тако, онда претпостављам да имамо тешкоће у разумевању термодинамичке ентропије. Из уобичајеног схватања „ентропије“ као „нереда“, што нам долази посматрајући чашу која падне са стола и разбије се правећи неред по поду, уз повећање своје ентропије, долази нам погрешна идеја да порастом ентропије расте информација датог термодинамичког система.

Претпоставимо ли да је идеја о расту информације заједно са ентропијом, нашли бисмо закључак да се термодинамички системи спонтано развијају у стања веће информације, а то би противречило начелном минимализму информације, принципу дејства физике (информација је еквивалентна дејству), па и очекивању да су вероватнији исходи чешћи (вероватније је мање информативно).

Дочим, порастом ентропије гас у посуди се хлади, његове молекуле мање осцилују, садрже мање неизвесности у свом кретању, мање информације. Ако вам ово није довољно убедљиво, приметимо и да те молекуле, са већом ентропијом, постају равномерније уређене. Попут војника постројених на смотри, оне постају безличне, мање информације нам емитују о себи, заправо, зато што су мање информативне – рећи ћемо сада – зато су једнолично уређене, да би информацију коју су исијали и изгубили надокнадили равномерношћу.

Овде се даље укључује горе поменута теорема о већој информацији једноликог распоређивања, сада због колико-толико одржања количине. Али уместо тога, приметимо да „говор делфина“ можемо препоставити на основу дисперзије сигнала, звукова које „размењују“, чак и не знајући о чему они „причају“. Уопште, пренос смислених порука, језик непознатих створења, наше текстове или разговор уопште карактеришу изразите неједнакости расподеле различитих слова, речи и уопште сигнала. То је зато што је различитост неопходна за разговетност, а тек потоњу има смисла мерити, или конзервирати њену количину.

10. Стохастичке матрице

Посебна врста матрице канала је стохастичка матрица. То је квадратна матрица, $\hat{S} = [\![s_{ij}]\!]$, реда n и ненегативних реалних коефицијената ($s_{ij} \geq 0$, за све $i, j = 1, 2, \dots, n$) чији се сваки ред, или свака колона, састоји од вероватноћа неких расподела. Ако су расподеле вероватноћа у врстама, хоризонталне, називамо је „десном“ због множења вертикалним вектором (контраваријантним) здесна и јединичног збира сваког ретка ($\sum_j s_{ij} = 1$).

Ако су расподеле вероватноћа у ступцима, вертикалне, стохастичку матрицу називамо „левом“ због множења хоризонталним вектором (коваријантним) слева и јединичног збира сваке колоне ($\sum_i s_{ij} = 1$). Стохастичка матрица је „двострука“, ако је квадратна матрица ненегативних реалних бројева са сваким редом и колоном јединичног збира.

Разматрајући адаптације каналом [08.] приметили смо да регуларна (детерминанта јој није нула) матрица канала има инверзну само ако су све њене врсте (или сви ступци) пермутације јединичне матрице. То се може строжије исказати и доказати.

Теорема 1. Стохастичка матрица $\hat{A} = [\![a_{ij}]\!]$ реда n има инверзну стохастичку матрицу $\hat{B} = [\![b_{ij}]\!]$ ако и само ако је добијена пермутацијама колона (редова) јединичне матрице $\hat{I} = [\![\delta_{ij}]\!]$, где је $\delta_{ij} = 1$ када је $i = j$, а ако $i \neq j$ онда је $\delta_{ij} = 0$. Инверзна матрица ($\hat{B}$) тада је транспонована дата ($\hat{A}$), дакле $\hat{B} = \hat{A}^{-1} = \hat{A}^{\tau}$, тј. $\hat{A}\hat{A}^{\tau} = \hat{A}^{\tau}\hat{A} = \hat{I}$.

Доказ: Претпоставимо да постоји $\hat{B}$ такво да је $\hat{B}\hat{A} = \hat{I}$, тј.

$$\sum_{k=1}^{n} b_{ik} a_{kj} = \delta_{ij}$$

По претпоставци теореме су матрице $\hat{A}$ и $\hat{B}$ стохастичке, што значи да су им сви коефицијенти ненегативни, па је сваки поједини (за свако $k = 1, 2, \dots, n$) сабирак $b_{ik} a_{kj} = 0$ за $i \neq j$. Како је $\hat{B}$ стохастичка, имаће у сваком ретку бар један коефицијент различити од нуле. Нека је то i-ти ред и $b_{ik} \neq 0$, па поделимо горњу једнакост овим коефицијентом. Добијамо

$$a_{kj} = \begin{cases} 1, & j \neq i \\ 0, & j = i \end{cases}$$

где је $a_{kj} = 1$, јер је матрица $\hat{A}$ стохастичка. Према томе, свака врста ове матрице има по једну јединицу и све остале коефицијенте нула.

Ових јединица матрица $\hat{A}$ има тачно n. Када би нека њена колона имала све коефицијенте нуле, онда производ $\hat{B}\hat{A}$ не би дао јединичну матрицу. Зато и свака колона матрице $\hat{A}$ има по једну јединицу. Према томе, она је добијена пермутацијом колона (врста) јединичне матрице.

Коначно, полазећи од горње једнакости, за $i = j$ добијамо $\sum_m b_{im} a_{mj} = 1$, те $b_{ik} a_{ki} = 1$ за свако поједино k, јер су сви остали сабирци (за $m \neq k$) нула. Отуда једнакост $a_{ki} = c$ даје $b_{ik} = c$, где је c константа 0 или 1, што је по дефиницији транспоновање. Тиме смо доказали и да је матрица $\hat{B}$ транспонована $\hat{A}$, те да су обе добијене пермутацијама колона јединичне матрице. ■

Даље се може доказивати да свака матрица канала $\hat{K} = [\![k_{ij}]\!]$ настала пермутацијама јединичне, има инверзну стохастичку матрицу.

Наиме, нека је $k_{ij} = 1$. Тада су сви остали елементи врсте и колоне нуле. Транспоновањем, тај се елеменат, i-те врсте и j-те колоне, нађе у i-тој колони и j-тој врсти. Множећи $\hat{K}\hat{K}^{\tau}$, i-ту врсту прве са i-том колоном друге добијамо $1 \cdot 1 = 1$, док у свим осталим колонама друге матрице на j-том месту имамо $1 \cdot 0 = 0$.

Упоређујући ово са осталим матрицама канала [01.], приметимо да у случају неквадратне матрице $\hat{K}$ из претпоставке $\hat{K}'\hat{K} = \hat{I}$, као у теореми 1, не може извести закључак да су матрице $\hat{K}$ и $\hat{K}'$ добијене пермутацијама колона јединичне матрице. То се јасније види из следећег примера.

Пример 1. Производ неквадратних матрица канала је јединична матрица ($\hat{K}'\hat{K} = \hat{I}$):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0{,}3 & 0{,}7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

али оне саме нису пермутације јединичне. ❑

Пример 2. Постоје две стохастичке матрице другог реда које су саме себи инверзне:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Оне су и симетричне. То су једине квадратне матрице другог реда које имају инверзну стохастичку матрицу. ❑

Пример 3. Постоји $3! = 3 \cdot 2 \cdot 1 = 6$ матрица реда $n = 3$ које имају инверзну стохастичку матрицу:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

рецимо означене редом $\hat{S}_k$. Последње две нису симетричне, па је нпр. $\hat{S}_5\hat{S}_6 = \hat{I}$. ❑

Многе теореме и примере узимам овде из моје књиге „Математичка теорија информације и комуникације'', у издању Друштва математичара Републике Српске, Бања Лука 1995. године. Немам је у дигиталном издању (писана је раних 80-их 20. века, али је због мог оклевања, дугих рецензија, а на крају и грађанског рата штампана 15-ак година касније), па користим прилике да понешто из те књиге извучем на интернет.

Због прегустог математичког стила (дефиниција, теорема, доказа, штурих примера) та је књига прошла незапажено, нечитано, а мислим да би се исто поновило у случају репринта, али ми је интересантно овде упоређивати неке тадашње ставке са новима, те сада већ „класичне теорије“ информације и комуникације са новима, мојом садашњом.

11. Адаптација стохастичком матрицом

Видели смо [10.Теорема 1] да стохастичка матрица да би имала инверзну мора у сваком ретку и колони имати по једну јединицу и све остале нуле. Таква је пермутација јединичне и делујући на вектор расподеле само пермутира (замењује места) његових коефицијената. Она тиме не мења дисперзије [09.] улазних расподела и не врши адаптацију вектора које пресликава. Разлогу [08.] адаптације вектора матрицом додајемо још једну, следећу теорему.

Теорема 1. Нека су $\vec{x} = (x_1, ..., x_n)$ и $\vec{y} = (\vec{y}_1, ..., \vec{y}_n)$ вектори две произвољне, али такве расподеле да је $\vec{y} = \hat{S}\vec{x}$, где је $\hat{S} = [\![s_{ij}]\!]$ дупла стохастичка матрица са сопственим вектором $\vec{p} = (p_1, ..., p_n)$. Тада је

$$\max_{1\le k\le n} |p_k - y_k| \le \max_{1\le j\le n} |p_j - x_j|$$

где важи једнакост ако и само ако из $p_i - x_i < \max_j |p_j - x_j|$ следи $s_{ij} = 0$ за свако $i = 1, ..., n$.

Доказ: Одузимањем добијамо $\vec{p} - \vec{y} = \hat{S}(\vec{p} - \vec{x})$. У том систему уочимо m-ту једначину где је апсолутна разлика $|p_m - y_m|$ максимална, тј. једначину $p_m - y_m = \sum_k (p_k - x_k)c_{kj}$. Како је $\sum_k c_{kj} = 1$ добијамо даље:

$$\max_k |p_k - y_k| = |p_m - y_m| \le \sum_k c_{ik}|p_k - x_k| \le \max_k |p_k - x_k|$$

што је и требало доказати. ∎

То је скоро исти претходни доказ овде [08. Адаптације канала], или у блогу (Minimalism), али и његове варијације вреде. Све овде речено о матрицама канала, чији су коефицијенти реални бројеви између 0 и 1, лако се преноси на комплексне матрице. Ставови тада важе за модуле коефицијената, односно квадрате модула (рецимо, квантно механичких интерпретација).

Оно што смо већ добили, или бисмо генерализацијама на комплексне бројеве могли добити, говори о „принципу минимализма информације“ језиком алгебре матрица. Како вектори (стања) тако и њихови линеарни оператори (процеси) конвергирају расподелама „мирнијих“ дисперзија, мање расутих вредности вероватноћа исхода, кажемо „адаптирају“ се.

Својствени вектори су те „тачке нагомилавања“ пресликавања, па ћемо допунити [06.] о њима. Већ смо видели да матрица канала $\hat{K} = [\![k_{ij}]\!]$ која допушта бар једну грешку ($0 < k_{ij} < 1$) нема инверзну матрицу. У композицији $(\hat{K} \cdot \hat{K} \cdots \hat{K})\vec{x} = \hat{K}^n\vec{x} \to \lambda\vec{p}$, када $n \to \infty$, где је $\vec{p}$ својствени вектор матрице $\hat{K}$ који одговара својственој вредности λ тако да је $\hat{K}\vec{p} = \lambda\vec{p}$.

Интуитивно можемо разумети да губимо податке у каналу ако бар један коефицијент није 0 нити је 1, али то је само грубо „разумевање“, недовољно и заправо нетачно. Наиме, детерминанта те матрице не мора бити нула (Channel), што је алгебарски довољан услов за инвертибилност. Матрице „канала са сметњом“, коефицијентом различитим од 0 и од 1, памте оригинале, али не дозвољавају пресликавање уназад. Оне су репрезентације једносмерних процеса: спонтаног раста ентропије, развоја система у мање информативне, иначе временског тока.

Матрице канала са сметњом дозвољавају дешифровање улаза на основу излаза, али не и стварно враћање копије назад у оригинал. Оне дозвољавају „псеудо инверзију“ и третирају саме процесе

слично стањима реалности и псеудо реалности (Dimensions), односно исказима који су тачни и нетачни (The Truth). За разлику од општих сопствених величина [06.], другачија ограничења стохастичких матрица дају и другачије конкретније резултате.

Теорема 2. Својствена једначина двоструке стохастичке матрице има само једну својствену вредност.

Доказ: Из својствене једначине $\hat{S}\vec{p} = \lambda\vec{p}$ добијамо:

$$\begin{cases} s_{11}p_1 + \cdots + s_{1n}p_n = \lambda p_1 \\ \dots \\ s_{n1}p_1 + \cdots + s_{nn}p_n = \lambda p_n \end{cases}$$

Сабирањем свих једначина, због $\sum_{k=1}^{n} s_{kj} = 1$ за свако $j = 1, \dots, n$, има само једну својствену вредност $\lambda = 1$. ∎

Знајући ову својствену вредност двоструке стохастичке матрице, на уобичајене начине тражимо својствене векторе.

Пример 1. За двоструку стохастичку матрицу $\hat{S}$ другог реда тражимо својствени вектор, бинарну расподелу $\vec{p} = (p_1, p_2)$, где је $p_1 + p_2 = 1$, а $p_1, p_1 \geq 0$, тако да је:

$$\hat{S}\vec{p} = \vec{p}$$

$$\begin{bmatrix} s & 1-a \\ 1-s & s \end{bmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_1 \end{pmatrix}$$

$$\begin{cases} sp_1 + (1-s)p_2 = p_1 \\ (1-s)p_1 + sp_2 = p_2 \end{cases}$$

$$p_1 = p_2 = \tfrac{1}{2}.$$

Према томе, својствени вектор је попут бацања фер новчића. ❑

Пример потврђује наша очекивања код спонтаног раста ентропије, када опада информација датог система чиме он постаје безличан. Спонтани раст ентропије датог система резултат је начелног минимализма комуникације (принципа најмањег дејства), а компромис са законом одржања информације уједначава вероватноће исхода. То смо већ видели у (новој) теорији информације, као поопштење горњег примера на на n-члане векторе, расподеле са укупно n дисјунктних исхода. Следећи став можемо сматрати на тај начин доказаним.

Теорема 3. Вектор $\vec{p} = (p_1, \dots, p_n)$ је својствени вектор дупле стохастичке матрице $\hat{S} = [\![s_{ij}]\!]$ ако и само ако $(p_i - p_j)s_{ij} = 0$ за све $i, j = 1, \dots, n$. ∎

Чисто алгебарски доказ овог става откријте сами, или га потражите у поменутој књизи „Математичка теорија информације и комуникације“.

12. Екстреми

Када мало боље погледамо, поменуте форме често се своде на два вектора, низа, $a = (a_1, \dots, a_n)$ и $b = (b_1, \dots, b_2)$, са скаларним производом $S = a \cdot b = a_1 b_1 + \dots + a_n b_n$. За сабирање важи закон комутације, па увек можемо преместити ове сабирке и пренумерисати низове тако да је рецимо први мнонотоно нерастући, тј. $a_1 \ge a_2 \ge \dots \ge a_n$. То не умањује општост наредних разматрања.

Деси ли се да је и други тада нерастући $b_1 \ge b_2 \ge \dots \ge b_n$, онда ће скаларни производ $S = a \cdot b$ бити максималан (M) у односу на било који поредак коефицијената низа b, уз исти поредак низа a. У обрнутом случају, када су ови низови супротне монотоности, први растући а други опадајући, онда је збир производа њихових одговарајућих парова минималан (m). Доследно овоме, било како пермутирали чланове тих низова и пренумерисали им индексе, добијати ћемо збир парова одговарајућих фактора $S = a_1 b_1 + \dots + a_n b_n$ који је увек ограничен, $m \le S \le M$, тада константним поменутим екстремима ($m \le M$).

Из ненегативних прираштаја $\Delta a = a_1 - a_2 \ge 0$ и $\Delta b = b_1 - b_2 \ge 0$, када је $\Delta a\, \Delta b \ge 0$, изводимо неједнакост $a_1 b_1 + a_2 b_2 \ge a_1 b_2 + a_2 b_1$. То је један од доказа (Emergence II) да се пермутовањем фактора двају сабирака њихов збир увећа, ако добијемо да се већи множи већим а мањи мањим. Сваким таквим кораком, унутар дужег низа сабирака S такође, повећава се вредност збира за неко $\Delta a\, \Delta b$. Ове замене редоследа вероватноћа слабије утичу на дисперзије расподела a и b, али и оне мењају вредност информације перцепције $S = a \cdot b$.

Другачији начин да приметимо исте промене били су скаларни производи вектора. Тада је такође $S = |a||b| \cos \varphi$, где је сада $\varphi = \sphericalangle(a, b)$ угао између два вектора. Интензитети вектора, $|a|$ и $|b|$, пермутацијама компоненти не мењају се, већ се мења само тај угао φ који они заклапају. При смањивању угла (ка нули), косинус расте. Када је угао φ све мањи, расте скаларни производ, два су вектора све ближа, боље „адаптирана“ један другоме, постају исте монотоности.

Аналогно скаларном производу вектора ($a \cdot b = |a||b| \cos \varphi$), уводили смо и „псеудо скаларни“ производ вектора a и b, помоћу чије вредности ($|a \times b| = |a||b| \sin \varphi$) смо извели формулу за ротацију вектора [04. Пример 1]. Када се први од та два производа повећава, други се смањује и обрнуто. Они се понашају попут вероватноће и информације, као две обрнуто сродне величине. Познато је да већој вероватноћи одговара мања информација. На пример, када знамо да је догађај известан, да ће се сигурно десити, а такав се деси, онда он и није нека вест.

Свеједно о којој се ради, било промени вероватноће или информације, имамо насилну појаву. Силе су те које мењају вероватноће – у (мојој) теорији информације. За саме пермутације исхода потребне су такве „силе“, па и регуларна дупла стохастичка матрица [10. Теорема 1] представља неки напор. Из [10. Пример 2] следи да је идентична матрица таква, те је идентично пресликавање врста, можда минималног, напора. Отуда идеја теорије информације да је сваки процес неки ток информација и дејство, да је сама информација еквивалент дејству. А тај смо закључак изводили и на друге начине (Action).

Како је физичко дејство производ размењене енергије и протеклог времена, или импулса и положаја, то производња информације, па и само копирање, ма каквим рачунарима, нису могући без трошења енергије. Не због самих машина које енергија покреће, већ и зато што произашла

информација није попут духа, без енергије, импулса, места или трајања, већ је информација еквивалентна физичком дејству. Уз везу информације и дејства долази и закон одржања.

Како је и само присуство информације неки „напор“ (сетимо се, стања и процеси су дуалне појаве), физички системи тежиће ка мање информативним стањима (поред чињенице да су та вероватнија и, према томе, изгледнија). Са друге стране, тој се тежњи супротставља закон одржања, па оне улазе у клинч и информација цури тамо где може. Не да се уништити попут енергије. Системе које је могуће посматрати у једнаким временским интервалима (а углавном су такви), не видимо као информативне већ као енергетске. Како год, оба имају законе одржања.

Даља последица закона одржања информације (па тако и енергије) је наизменично смењивање привлачних и одбојних „сила“, како у промени скале величина, тако у ближим и даљим атомима или молекулама, у напрезању и отпуштању еластичних тела. Овде се исто очитује у компензацији, у жељи природе (метафорички) заустављања губитка информације приликом процеса спонтаног њеног опадања. Зато се хлађењем тела, при повећању ентропије и губитку информације његових молекула (нпр. мање неочекиваних њихових вибрација), вероватноће исхода смирују, дисперзије расподела се смањују – као да његова укупна информација настоји остати већом.

Трећа „сила“ која се овде појављује је тежња ка многоструккошћу. Једва је приметна и потцењена, али је веома значајна. То је већ расправљано у наслову [09.], или у блогу (Channel, Теорема). Ако су два ретка (две колоне) неке квадратне матрице $\hat{M}$ пропорционалне, детерминанта јој је нула ($\det \hat{M} = 0$), систем линеарних једначина које она представља ($\vec{y} = \hat{M}\vec{x}$) нерегуларан је, што значи да нема јединственог решења или нема решења уопште, те да нема ни инверзног пресликавања ($\vec{x} = \hat{M}^{-1}\vec{y}$). Слично ће важити и за приближно пропорционалне редове матрице. То је садржај поменуте теореме.

Како су редови матрице канала расподеле вероватноћа, они безлични (уједначених вероватноћа) неће преносити поруке. Из њихових копија ($\vec{y}$) није могуће дешифровање ($\vec{x}$), нити инверзно пресликавање (могуће је дешифровање без инверзног пресликавања [11.]). Последице дисперзије (различитости) вероватноћа су: инвертибилност процеса, затим могућност дешифровања поруке на основу пријема, па и закон одржања. Смањењем безличности, дакле појавом различитости прелази се праг након којег је то препознавање порекла све јасније до праве реверзибилности.

Поред познатих дисперзија статистике, горњи екстреми ($m \le S \le M$) су мера њених граница. Број $|M - m|$ зависи од дисперзије вероватноћа, од броја могућих исхода, па и од ефикасности система којег представља. Ово последње видимо на моделу „слободних мрежа“. Ако су повезнице једнако вероватне, онда се мрежа развија ка малом броју чворова са много повезница наспрам много њих са мало повезница, попут слободног тржишта роба и новца са власницима у чворовима. То повећава ефикасност тока информација слободне мреже на уштрб равноправности чворова.

13. Многострукост

У претходном наслову поменуто је да постоје три врсте спонтаних „сила“ које покрећу свет информација: минимализам, одржање и различитост. Зато што би се случајни исходи радије реализовали у вероватније, они би ишли у мање информативне, али уз одржање количине неизвесности. Сва та количина пре бацања коцке ($\log 6$) претвара се у исход након, али у многим другим случајевима имамо задршку. Често само један део неизвесности одлази у исход, тако да понешто од њега остаје као неизјашњена информација.

Знајући да важи закон одржања информације ове „задршке“ умемо препознавати. Искакање честица из празног простора (вакуума) је једна таква појава, друга је настајање својстава ентитета којих његови делови сами по себи немају (Emergence II), трећа је појава ефикасности „слободних мрежа“, а има их неограничено.

Пример 1. Матрица канала, система $\vec{b} = \hat{M}\vec{a}$, не мора бити стохастичка

$$\begin{pmatrix} 0,73 \\ 0,23 \\ 0,04 \end{pmatrix} = \begin{bmatrix} 0,73 & 0,73 \\ 0,23 & 0,23 \\ 0,04 & 0,04 \end{bmatrix} \begin{pmatrix} 0,5 \\ 0,5 \end{pmatrix}$$

иако вектори $\vec{a}$ и $\vec{b}$ представљају расподеле. Њихове Шенонове информације су:

$$\mu(\vec{a}) = -a_1 \log_2 a_1 - a_2 \log_2 a_2 = -0,5 \cdot \log_2 0,5 - 0,5 \cdot \log_2 0,5 = 1$$

$$\mu(\vec{b}) = -b_1 \log_2 b_1 - b_2 \log_2 b_2 - b_3 \log_2 b_3 =$$

$$= -0,73 \cdot \log_2 0,73 - 0,23 \cdot \log_2 0,23 - 0,04 \cdot \log_2 0,04 \approx 1.$$

Дакле, $\mu(\vec{a}) = \mu(\vec{b})$, информације овако различитих расподела могу бити једнаке. ❑

Пример демонстрира могућност појаве „неочекиваности“ ради одржања количине (информације) уз развој различитости. Приметимо да улазна 2-компонентна расподела $\vec{a}$ има највећу средњу информацију (Шенонову) коју таква може имати, а да излазна $\vec{b}$, да би била 3-компонентна и да би очувала укупну информацију, постаје израз веће ефикасности, тј. мање замућености исхода. Средња вредност μ јој остаје иста, али јој је дисперзија је већа. Матрица $\hat{M}$, са истим учинком, може имати и друге коефицијенте.

Просечна информација система, попут Шенонове, мера је слична специфичној густини супстанце и као таква има своју улогу у (мојој) теорији информације, поред Физичке информације. Одржавање ове „специфичне информације“, каква је μ у примеру, могуће је постићи и растом многострукости, али по цену веће опредељености. Иначе ће, понављам, уједначавањем вероватноћа она порасти.

Теорема 1. Ако за две вероватноће неке расподеле важи неједнакост $p_1 < p_2$ тада је

$$\mu(p_1 + \varepsilon, p_2 - \varepsilon, p_3, \dots) > \mu(p_1, p_2, p_3, \dots)$$

за довољно мало $\varepsilon > 0$.

Доказ: Потражимо инфинитезималну промену средње вредности информације, када се прва од вероватноћа расподеле повећа, а друга смањи. Тотални диференцијал је

$$d\mu(\vec{p}) = \sum_k \frac{\partial \mu}{\partial p_k} dp_k$$

где је $dp_1 = \varepsilon, dp_2 = -\varepsilon, dp_3 = 0, \dots$, а $\frac{\partial \mu}{\partial p_k} = -\log p_k - \log e$, то је $d\mu = \varepsilon \cdot \log \frac{p_2}{p_1} > 0$, јер је $\frac{p_2}{p_1} > 1$, што је и требало доказати. ■

То је сасвим очекиван резултат, да уједначавањем вероватноћа расподеле, средња вредност информације (Шенонове) расте. Обрнуто, нарушавањем једнакости вероватноћа, информација опада, па је за одржавање њене константности потребно повећати број исхода расподеле, као што је то учињено у примеру 1. Ова теорема је, дакле, поопштење примера попут наведеног.

Пример 2. Систем једначина $\vec{b} = \hat{M}\vec{a}$ је

$$\begin{pmatrix} 0,75 \\ 0,20 \\ 0,05 \end{pmatrix} = \begin{bmatrix} 1,36 & 0 \\ 0 & 0,44 \\ 0,05 & 0,05 \end{bmatrix} \begin{pmatrix} 0,55 \\ 0,45 \end{pmatrix}$$

где су вектори $\vec{a}$ и $\vec{b}$ расподеле. Њихове средње информације су:

$$\mu(\vec{a}) = -0,55 \cdot \log_2 0,55 - 0,45 \cdot \log_2 0,45 = 0,99$$

$$\mu(\vec{b}) = -0,75 \cdot \log_2 0,75 - 0,20 \cdot \log_2 0,20 - 0,05 \cdot \log_2 0,05 = 0,99$$

дакле (приближно) једнаке. ❑

Ово је такође пример „бежања у многострукост“ да би се избегла равномерност. Расподела $\vec{a}$ добија трећу компоненту (0,05) да би повећала разлику $(0,55 - 0,45) \rightarrow (0,75 - 0,20)$ водећих вероватноћа. Међутим, ако је та почетна разлика мало већа, рецимо $\vec{a} = (0,6; 0,4)$ или (0,7; 0,3), или још већа, онда се овакво „бежање“ не може десити. Природа не воли једнакости, али неће претеривати.

Када не претерује са истом методом, природа може „из рукава“ извући нову методу и опет нас изненадити својом многострукошћу. Методе као и процеси, или стања, подложни су сличним формама и „чудима“. Пример таквог једног изненађења је делимични одлазак сопственог посматрача у псеудо реалност, у односу на релативног, у јачем гравитационом пољу, или при релативном кретању.

Непредвидљиост је објективна појава за сваког појединог учесника овог света (честице, тела, нас). Она је до те мере реална да је ипак спознатљива, уз одређене напоре (промене вероватноћа). То делује парадоксално, али да није контрадикторно сведочи пример ловца који би хватао дивљач својом лукавошћу и клопком, која у датом тренутку није спознатљива његовом плену. Слојевитост и релативности информације подлога су објективности њој потребне неизвесности.

14. Хипер матрица канала

Матрица канала $\hat{K} = [\![k_{ij}]\!]$ типа $m \times n$ има ненегативне коефицијенте, $k_{ij} \geq 0$, распоређене у m редова (индекса i) и n колона (j). Они не морају бити расподеле вероватноће, па ни њихов укупни збир не мора бити Шенонова информација ($-\sum_i \sum_j k_{ij} \log k_{ij}$), нити нека друга средња вредност.

Нека је дат такав систем једначина $\hat{K}\vec{p} = \vec{q}$, или детаљније:

$$\begin{cases} k_{11}p_1 + \cdots + k_{1n}p_n = q_1 \\ \cdots \\ k_{m1}p_1 + \cdots + k_{mn}p_n = q_m \end{cases}$$

Он дефинише једну линеарну трансформацију вероватноћа расподеле $\vec{p}$ у расподелу $\vec{q}$. Обратимо пажњу на ових mn сабирака $h_{ij} = k_{ij}p_j$ и формирајмо њихову матрицу

$$\hat{H} = [\![h_{ij}]\!] = \begin{bmatrix} k_{11}p_1 & \dots & k_{1n}p_n \\ \dots & \dots & \dots \\ k_{m1}p_1 & \dots & k_{mn}p_n \end{bmatrix}.$$

Назовимо је „хипер матрицом канала“. Она представља оба вектора, па има и ознаку $\hat{H}(\vec{p},\vec{q})$.

Сваки, i-ти, ред матрице $\hat{K}$ неки је низ, вектор $\vec{K}_{i*}$, а аналогно је и са колонама $\vec{K}_{*j}$. Када овакав чини неку расподелу вероватноћа можемо говорити о средњим вредностима информација ретка и колоне:

$$\mu(\vec{K}_{i*}) = K_{i*} = -\sum_{j=1}^{n} k_{ij} \log k_{ij}, \quad \mu(\vec{K}_{*j}) = K_{*j} = -\sum_{i=1}^{m} k_{ij} \log k_{ij}.$$

Оне су тада заправо Шенонове информације. Међутим

$$\mu(\hat{K}) = \sum_{i=1}^{m} K_{i*} = -\sum_{i,j}^{m,n} k_{ij} \log k_{ij}$$

није у правом смислу средња вредност, тј. није Шенонова информација. Али следећа јесте Шенонова информација (јер h_{ij} јесте расподела):

$$\mu(h_{ij}) = -\sum_{i=1}^{m}\sum_{j=1}^{n} h_{ij} \log h_{ij}$$

Доследно претходном можемо писати $\mu(h_{ij}) = \mu(\vec{p},\vec{q})$. То је средња вредност информације самог овде названог „хипер канала“ коју треба разликовати од улазних и излазних:

$$\mu(\vec{p}) = -\sum_{j=1}^{n} p_j \log p_j, \quad \mu(\vec{q}) = -\sum_{i=1}^{m} q_i \log q_i = -\sum_{i=1}^{m}\sum_{j=1}^{n} h_{ij} \log \sum_{l=1}^{n} h_{il}$$

средњих вредности информација оригинала и копије.

Када је K_{i*} Шенонова информација i-тог ретка (неизвесност улаза $\vec{p}$ када је извесан излаз $\vec{q}$), а q_i њена вероватноћа, или аналогно обрнуто, тада су:

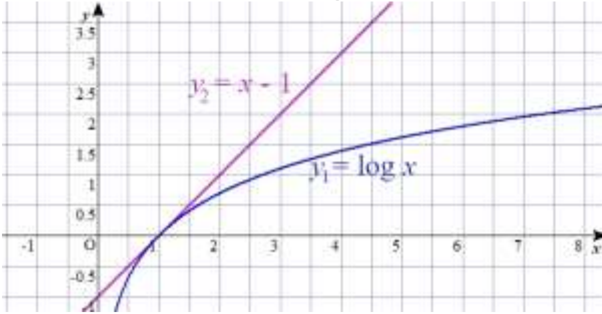
$$\mu(\vec{p}|\vec{q}) = \sum_{i=1}^{m} q_i K_{i*}, \quad \mu(\vec{q}|\vec{p}) = \sum_{j=1}^{n} p_j K_{*j}$$

средње вредности информација редова и колона канала $\widehat{K}$.

Сви коефицијенти хипер матрице канала ненегативни су ($h_{ij} = k_{ij}p_j \ge 0$), а њихов збир је јединица ($\sum_{i,j} h_{ij} = 1$), па они увек представљају неку расподелу вероватноћа. Како већој шанси (h_{ij}) одговара мања информација ($-\log h_{ij}$) и обрнуто, то је „$-\sum_{ij} h_{ij} \log h_{ij}$“ скаларни прозвод два низа од по „$m \cdot n$“ чланова. Други релативни поредак једног од таквих низова увећаће скаларни производ, ако се наруши правило „мањи са већим“ [12. Екстреми]. То значи да ће и хипер матрица канала других коефицијената h'_{ij} комбинована са овом дати већу вредност пратећег скаларног производа. Даље следи понављање градива, па то водимо као примере.

Пример 1. У прилогу (Emergence II, Теорема) је један локални доказ тог става у којем се користи „очигледна“ неједнакост $\log x \le x - 1$, односно ордината $y_1 \le y_2$, која се види на графу, где се види и да једнакост важи ако и само ако је апсциса $x = 1$. Наиме, сменама $x = h'_{ij}/h_{ij}$ за $h_{ij} \ne 0$ и сабирањем добијамо

$$-\sum_{ij} h_{ij} \log h_{ij} \le -\sum_{ij} h_{ij} \log h'_{ij}.$$



То нам говори да ће за ове матрице вредети „принцип минимализма“ информације, или ако нам је згодније рећи „принцип најмањег дејства“ теоријске физике, односно да чешће налазимо вероватније догађаје. ❑

Са „погрешном“ расподелом информација се повећава. Њен принципијелни минимализам стога је разлог удруживања, од квантног спрезања до формирање друштвених организација. Алгебра ове теорије је далеког домета.

Пример 2. Једна од „погрешних“ информација била би и K_{*j} схваћена као средња вредност неке (замишљење) просечне равномерне расподеле, када је комбинујемо са вероватноћама дате

$$-\sum_{j=1}^{n} p_j \log p_j \le -\sum_{j=1}^{n} p_j K_{*j}$$

а отуда $\mu(\vec{p}) \le \mu(\vec{p}|\vec{q})$. Аналогно $\mu(\vec{q}) \le \mu(\vec{q}|\vec{p})$. ❑

Пример 3. Средња вредност информације хипер матрице канала је:

$$\mu(h_{ij}) = -\sum_{i,j=1}^{m,n} h_{ij} \log h_{ij} = -\sum_{i,j=1}^{m,n} h_{ij} \log(k_{ij}p_j) =$$

$$= -\sum_{i,j=1}^{m,n} k_{ij}p_j(\log k_{ij} + \log p_j) = -\sum_{i,j=1}^{m,n} (p_j k_{ij} \log k_{ij} + k_{ij}p_j \log p_j)$$

$$= \sum_{j=1}^{n} (K_{*j}p_j - p_j \log p_j) = \sum_{j=1}^{n} K_{*j}p_j - \sum_{j=1}^{n} p_j \log p_j$$

$$= \mu(\vec{q}|\vec{p}) + \mu(\vec{p}),$$

тј. $\mu(\vec{p},\vec{q}) = \mu(\vec{p}) + \mu(\vec{q}|\vec{p})$. Слично се доказује $\mu(\vec{p},\vec{q}) = \mu(\vec{q}) + \mu(\vec{p}|\vec{q})$. ❑

Саре ознаке $H(*)$ треба читати као „ентропије канала...“, а овдашње $\mu(*)$ су „средње вредности информације...“. У новој „теорији информације“ (пишем наводној, која је још увек моја приватна) са спонтаним повећањем термодинамичке ентропије долази до смањивања информације, за разлику од тадашње и садашње физике у којој се понекад верује да су те величине директне.

Дефиниција 1. Начин за помирење нових и старих тумачења и ознака могао би бити у промени предзнака при преласку са информације на ентропију и обрнуто. Тако је:

$$H(\vec{p},\vec{q}) = -\mu(\vec{p},\vec{q}),\quad H(\vec{p}) = -\mu(\vec{p}),\quad H(\vec{q}) = -\mu(q),$$

$$H(\vec{p}|\vec{q}) = -\mu(\vec{p}|\vec{q}),\quad H(\vec{q}|\vec{p}) = -\mu(\vec{q}|\vec{p}).$$

Овде су μ подразумеване средње вредности информације, осим ако је другачије речено. ▨

Ово је само предлог за усклађивање појмова и ознака ентропије и информације, на основу којега би се следећи пример могао подвести под иначе познату теорему класичне „Теорије информације и комуникације“ (в. поменуту моју истоимену књигу).

Пример 4. За сваки пар расподела $\vec{p}$ и $\vec{q}$ за ентропије важе релације:

$$H(\vec{p}) \ge H(\vec{p}|\vec{q}),\quad H(\vec{q}) \ge H(\vec{q}|\vec{p}),$$

$$H(\vec{p},\vec{q}) = H(\vec{p}) + H(\vec{q}|\vec{p}) = H(\vec{q}) + H(\vec{p}|\vec{q}),$$

$$H(\vec{p}) + H(\vec{q}) \ge H(\vec{p},\vec{q}).$$

Прва и друга неједнакост следе из примера 2 и дефиниције 1. Оне постају једнакости, због примера 1, ако и само ако је $q_j = k_{ij}$, што множећи са p_i даје $h_{ij} = p_i q_j$. Среде једнакости су непосредне последице примера 3 и дефиниције, апоследња неједнакост излази сабирањем прве две и средњих. ❑

На пример, дата је матрична једнакост $\vec{q} = \hat{K}\vec{p}$, односно систем једначина:

$$\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{bmatrix} 0{,}5 & 0{,}3 & 0{,}2 \\ 0{,}5 & 0{,}7 & 0{,}8 \end{bmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

$$\begin{cases} q_1 = 0{,}5p_1 + 0{,}3p_2 + 0{,}2p_3 \\ q_1 = 0{,}5p_1 + 0{,}7p_2 + 0{,}8p_3 \end{cases}$$

где је пресликавање $\hat{K}: \vec{p} \to \vec{q}$, рецимо $(0{,}6; 0{,}2; 0{,}2) \to (0{,}4; 0{,}6)$. Овде је $\sum_{i,j=1}^{2,3} k_{ij}p_j = 1$, па је ово хипер матрица канала. Она расподелу пресликава у расподелу, али та матрица није стохастичка, јер њен други ред није расподела вероватноћа. Збир коефицијената њеног ($\hat{K}$) другог ретка већи је од један, па је непримењив став примера 1. Овај пример, међутим, приказ је компликованости односа уведених појмова.

15. Канали комплексних бројева

Сматрамо ли теорију комплексних бројева тачном, они $\mathbb{C}$ су онда нека врста реалности и као такви део су теорије информације. Један од начина признавања комплексних бројева је „заобилажење“. Паралелне реалности нису „стварне“ већ „псеудо“, или рецимо „имагинарне“, али су неизбежне за комплетирање информатичог света, једнако као што су комплексни бројеви потребни алгебри. Покушајмо то разумети на примеру својствених вектора.

Матрица ротације [03.] је реална, има само реалне коефицијенте, пресликава реалне векторе (са реалним коефицијентима) у реалне, али она нема реалан својствени вектор. Сваком ће вектору променити смер, заротирати га, а не само га умножавати λ пута. Својствени вектор $\vec{x}$ пресликан матрицом $\hat{A}$ остаје на истом правцу, продужен или скраћен за неку својствену вредност, $\hat{A}\vec{x} = \lambda\vec{x}$.

Пример 1. Својствена једначина ротације, са комплексним векторима ($\hat{R}\vec{z} = \lambda\vec{z}$) даје:

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\begin{bmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$(\cos\varphi - \lambda)^2 + \sin^2\varphi = 0$$

$$\lambda_{1,2} = \cos\varphi \pm i\sin\varphi = e^{\pm i\varphi},$$

а ово није реално. За својствене вектор, сменом ламбди, добијамо $y = \pm ix$, а то су, на пример, јединични вектори $z_{1,2} = (1 \pm i)/\sqrt{2}$. ❑

Као што за комплексне бројеве који „побегну“ изван реалног система координата не сматрамо да су отишла у „свет контрадикција“, нетачности, тако ни за имагинарна, или комплексна решења физикалних ситуација нећемо рећи да су изван сваке реалности.

Што се тиче ротација, из примера видимо да не треба ићи даље од комплексних својствених вектора (оних који „трају“), а како се помоћу ротација могу добијати све остале изометријске трансформације (које чувају удаљености између тачака), то се „трајања“ кроз све физикално занимљиве процесе могу свести најдаље на комплексне бројеве.

Други начин да докажемо исто је помоћу основне теореме алгебре, да сваки полином има бар једну нулу, решење у скупу реалних или комплексних бројева, из чега следи да полином n-тог степена има n решења и да су сва најдаље комплексна. Затим приметимо да је карактеристична, својствена једначина $(\widehat{M} - \lambda\hat{I})\vec{x} = 0$, где је λ својствена вредност матрице $\widehat{M}$ а $\vec{x}$ одговарајући својствени вектор – заправо полиномска.

Томе додајмо да је (апстрактни) простор матричних репрезентација еквивалентан линеарним операторима, а да су подврсти ових (Хилбертовој алгебри) репрезентација квантна стања (вектори) и квантни процеси (унитарни оператори). Теорија информације овом трајању стања (вектора) кроз процесе (оператора) придружује шетње, тј. одласке и доласке из псеудо реалности у текућу садашњост.

Обзиром на ограничену брзину светлости, ни супротни крајеви иоле макро тела нису у истој садашњости. Особа која седи поред саговорника никада му није довољно близу да би њих двоје могли бити у тачно истој реалности, да би оно што чује било управо „сада“, тачно тада када је и изговорено. А „реалност“ прошлости врсте је псеудо реалности временских димензија.

Када имамо два својствена вектора, $\vec{x}_1$ и $\vec{x}_2$, неке својствене вредности λ датог оператора $\hat{A}$, онда је и њихова линеарна комбинација $\vec{x} = a_1\vec{x}_1 + a_2\vec{x}_2$ такође својствени вектро исте својствене вредности истог оператора. Наиме:

$$\hat{A}\vec{x} = \hat{A}(a_1\vec{x}_1 + a_2\vec{x}_2) = a_1\hat{A}(\vec{x}_1) + a_2\hat{A}(\vec{x}_2) = a_1\lambda\vec{x}_1 + a_2\lambda\vec{x}_2 = \lambda(a_1\vec{x}_1 + a_2\vec{x}_2) = \lambda\vec{x}.$$

То значи да својствени вектори разапињу потпросторе својствених вектора. Те векторске просторе називамо „својственим потпросторима“ матрице (као и оператора) $\hat{A}$.

Реалности и псеудо реалности, у оквиру којих наше садашњости „шетају“, чине један затворени систем „трајања“ процеса, у односу на дате операторе и њихове својствене вредности. Теоријом графова могуће их је приказивати аналогно удруживању јединки у колективе, или уопште у један спрегнути систем подлегао принципијелном минимализму информације.

У наставку погледајмо једну добро познату занимљиву теорему о комутативности оператора и инваријантности вектора. Подсећам, оператори су комутативни, ако њихово деловање не зависи од редоследа. Подпростор је инваријантан према линеарном оператору ако му и копије припадају подпростору.

Теорема 1. Када два комутативна оператора, или матрице $\hat{A}, \hat{B} \in (X \to X)$, припадају истом векторском простору, па ако је $X_0 \subseteq X$ својствени потпростор матрице $\hat{A}$ онда је он инваријантан с обзиром на $\hat{B}$.

Доказ: Из $\hat{A}\vec{x}_0 = \lambda\vec{x}_0$, следи $\hat{A}\left(\hat{B}x_0\right) = \hat{B}\left(\hat{A}\vec{x}_0\right) = \hat{B}(\lambda\vec{x}_0) = \lambda\hat{B}\vec{x}_0$, што значи да вектор $\hat{B}\vec{x}_0 \in X_0$, тј. X_0 је инваријантно с обзиром на $\hat{B}$. ∎

Примењено на процесе микро честица, комутативност значи сродност, спрегнутост. У теорији информације то је „истовременост“, подсећам, која ће се посматрачу који није истовремен са дотичним чинити као „фантомско деловање на даљину“, усклађивање без преноса информације. Са друге стране, Хајзенбергових релација, то значи да нема неодређености, $\left(\hat{A}\hat{B} - \hat{B}\hat{A}\right)\psi = 0$.

16. Инверзна матрица

Алгебра нас учи да систем n линеарних једначина са исто толико непознатих има једнаке скупове решења као и њему одговарајућа матрична једначина ($\hat{A}\vec{x} = \vec{y}$). Довољно је да детерминанта те матрице буде различита од нуле ($\det \hat{A} \neq 0$) да постоји инверзан систем једначина ($\vec{x} = \hat{A}^{-1}\vec{y}$) у којем се оригинали израчунавају помоћу копија. Производи таквих су јединична матрица:

$$\vec{x} = \hat{A}^{-1}\vec{y} = \hat{A}^{-1}\hat{A}\vec{x}, \quad \hat{A}^{-1}\hat{A} = \hat{I}: \vec{x} \to \vec{x},$$

$$\vec{y} = \hat{A}\vec{x} = \hat{A}\hat{A}^{-1}\vec{y}, \quad \hat{A}\hat{A}^{-1} = \hat{I}: \vec{y} \to \vec{y}\,.$$

Међутим, ако су низови оригинала и копија варијабли ($\vec{x}, \vec{y}$) расподеле вероватноћа, онда ће се десити нарушавање овог правила. Видели смо [11.] то у случају вредности детерминанти између 0 и 1, јер тада матрица канала сужава вредности улазних расподела, смањује дисперзију, које се враћањем назад више не могу поправити.

Размотримо речено полако на једноставном систему две једначине и две непознате, при чему ћемо дозволити комплексне бројеве за коефицијенте. Дат је систем $\hat{C}\vec{z} = \vec{c}$, односно:

$$\begin{cases} c_{11}z_1 + c_{12}z_2 = c_{13} \\ c_{21}z_1 + c_{22}z_2 = c_{23} \end{cases}$$

па прву множимо са c_{22} другу са c_{12} и одузмимо их да елиминишемо z_2. Ако бисмо прву множили са c_{21} а другу са c_{11} и одузели, елиминисали бисмо z_1. Тако налазимо:

$$\begin{cases} (c_{11}c_{22} - c_{21}c_{12})z_1 = c_{13}c_{22} - c_{23}c_{12} \\ (c_{11}c_{22} - c_{21}c_{12})z_2 = c_{11}c_{23} - c_{21}c_{13} \end{cases}$$

$$\begin{cases} \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} z_1 = \begin{vmatrix} c_{13} & c_{12} \\ c_{23} & c_{22} \end{vmatrix} \\ \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} z_2 = \begin{vmatrix} c_{11} & c_{13} \\ c_{21} & c_{23} \end{vmatrix} \end{cases}$$

$$\begin{cases} D \cdot z_1 = D_1 \\ D \cdot z_2 = D_2 \end{cases}$$

Ово је тзв. Крамеров систем из којег се види да у случају нујлте детерминанте система, $D = 0$, имамо пропорцију $c_{11}: c_{12} = c_{21}: c_{22}$, па ако важи и дужа пропорција $c_{11}: c_{12}: c_{13} = c_{21}: c_{22}: c_{23}$ систем је сагласан и има бесконачно решења, а иначе је у контрадикцији (нема решења). Када је детерминанте система различите од нуле, $D \neq 0$, постоји јединствено решење. То је овде пар бројева $z_k = D_k/D$, за $k = 1, 2$. Аналогно важи за више једначина и једнако томе непознатих.

Ставимо ли да су на десној страни једначина датог система варијабле $y_1 = c_{13}$ и $y_2 = c_{23}$, у случају јединственог решења, када је могуће дељење ($D \neq 0$) биће:

$$\begin{cases} z_1 = \dfrac{1}{D}(c_{22}y_1 - c_{12}y_2) \\ z_2 = \dfrac{1}{D}(c_{11}y_2 - c_{21}y_1) \end{cases}$$

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{1}{D}\begin{bmatrix} c_{22} & -c_{12} \\ -c_{21} & c_{11} \end{bmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\vec{z} = \hat{C}^{-1}\vec{y}$$

па је $\hat{C}^{-1}$ инверзна матрици $\hat{C}$. Заиста:

$$\hat{C}^{-1}\hat{C} = \frac{1}{D}\begin{bmatrix} c_{22} & -c_{12} \\ -c_{21} & c_{11} \end{bmatrix}\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} =$$

$$= \frac{1}{D}\begin{bmatrix} c_{22}c_{11} - c_{12}c_{21} & 0 \\ 0 & c_{22}c_{11} - c_{12}c_{21} \end{bmatrix}$$

$$= \frac{1}{D}\begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \hat{I}$$

а то је јединична матица (2×2). Слично проверавамо $\hat{C}\hat{C}^{-1} = \hat{I}$. Шта је онда тешко са инверзијом расподела, погледајмо на примеру квантно механичког тумачења вероватноће.

Из горњег система једначина, за k-ту компоненту варијабле имамо:

$$\|y_k\| = \left|\sum_j c_{kj} z_j\right| \leq \|\vec{c}_{k*}\|\|\vec{z}\| = c_k,$$

где је примењена Шварцова неједнакост, јер је расподела вероватноћа нормирана, $\|\vec{z}\| = 1$, а реална константа k-тог ретка матрице $c_k > 0$. Због $\|y_k\| \leq c_k$, ако је $c_k < 1$, онда излаз $(\vec{y})$ није расподела. А ако је $c_k > 1$, онда улаз $(\vec{z})$ није расподела (следи из инверзне матрице). Штавише, ако бар један од ових бројева за редове није један $(c_k \neq 1)$ немамо пресликавање расподела датом матрицом, иако она има инверзну матрицу.

Дакле, линеарна алгебра матрица канала у случају пресликавања расподела вероватноћа суочава се са додатним ограничењем. То је могућност егзистенције инверзне матрице, а да при томе није могуће инверзно пресликавање.

Овде смо разматрали само генерализацију „реалне“ (класичне) теорије вероватноће на „комплексну“, на начин квантне механике (Борнов закон). Приметите да резултати нису дословно пренешени. Тамо [11.] смо приметили ограничења када се редак (колона) не састоји од једне јединице и свих осталих нула, а овде када генералисани није јединичне норме.

На крају сетимо се да је детерминанта производа матрица једнака производу детерминанти тих матрица, затим да услов нормираности редака дате матрице значи и нормираност саме матрице. Тако долазимо до унитарних матрица, које су следећа тема.

17. Унитарна матрица

Матрица је унитарна када јој је детерминанта јединична, $\det \hat{U} = 1$. Ако је $\hat{U}$ унитарна, онда из $\det \hat{U} = 1$ коњуговањем[1] следи $\det \hat{U}^* = 1$, па је и њена коњугована матрица $\hat{U}^*$ унитарна.

Пример 1. Проверимо ово не примеру опште унитарне матрице другог реда:

$$\hat{U}\hat{U}^* = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix} = \begin{bmatrix} aa^* + bb^* & ac^* + bd^* \\ ca^* + db^* & cc^* + dd^* \end{bmatrix}$$

Како је $\det \hat{U} = 1$, то је $ad - bc = 1$, а отуда $a^*d^* - b^*c^* = 1$. Затим:

$$\det \hat{U}\hat{U}^* = (aa^* + bb^*)(cc^* + dd^*) - (ca^* + db^*)(ac^* + bd^*) =$$

$$= (aa^*cc^* + aa^*dd^* + bb^*cc^* + bb^*dd^*) - (ca^*ac^* + ca^*bd^* + db^*ac^* + db^*bd^*)$$

$$= aa^*dd^* + bb^*cc^* - ca^*bd^* - db^*ac^*$$

$$= (ada^*d^* - adb^*c^*) - (bca^*d^* - bcb^*c^*)$$

$$= ad(a^*d^* - b^*c^*) - bc(a^*d^* - b^*c^*)$$

$$= (ad - bc)(a^*d^* - b^*c^*) = 1.$$

Према томе $\det \hat{U}\hat{U}^* = 1$. Са друге стране, детерминанта производа матрица једнака је производу детерминанти тих матрица, и $\det \hat{U}^* = 1$. ❑

Замена редоследа множења даје слично $\det(\hat{U}^*\hat{U}) = 1$, а коњуговањем производа добијамо исти, $(\hat{U}^*\hat{U})^* = (\hat{U}^*)^*(\hat{U})^* = \hat{U}\hat{U}^*$. Отуда $\hat{U}^* = \hat{U}^{-1}$, тј. коњугована унитарна матрица је и иверзна.

Пример 2. Из једнакости:

$$\begin{cases} \hat{U}\hat{U}^* = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix} = \begin{bmatrix} aa^* + bb^* & ac^* + bd^* \\ ca^* + db^* & cc^* + dd^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \hat{U}^*\hat{U} = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^*a + c^*c & a^*b + c^*d \\ b^*a + d^*c & b^*b + d^*d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{cases}$$

следи:

$$\begin{cases} aa^* + bb^* = cc^* + dd^* = a^*a + c^*c = b^*b + d^*d = 1 \\ ac^* + bd^* = ca^* + db^* = a^*b + c^*d = b^*a + d^*c = 0 \end{cases}$$

Из прве четири, одузимањима налазимо $aa^* = d^*d$ и $cc^* = b^*b$, а затим да су ови дијагонални коефицијенти заротирани за неки угао и сва четири на јединичној кружници комплексне равни. Сменом у доње четири налазимо $d = a^*e^{i\varphi}$ и $c = -b^*e^{i\varphi}$, где је φ произвољан реалан број (угао) и $|a|^2 + |b|^2 = 1$. Према томе:

$$\hat{U} = \begin{bmatrix} a & b \\ -b^*e^{i\varphi} & a^*e^{i\varphi} \end{bmatrix}, \quad \hat{U}^* = \begin{bmatrix} a^* & -be^{-i\varphi} \\ b^* & ae^{-i\varphi} \end{bmatrix}$$

---

[1] Комплексан број је $z = x + iy$, где су x и y реални бројеви, а за имагинарну јединицу важи $i^2 = -1$. Њему коњугован број је $z^* = x - iy$.

$$\hat{U}\hat{U}^* = \hat{U}^*\hat{U} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}.$$

Дакле $\hat{U}^{-1} = \hat{U}^*$, ова унитарна матрица и њена коњугована међусобно су инверзне. ❑

Комплексна квадратна матрица реда $n = 1, 2, 3, \dots$ која комутира са коњугованом, $\hat{M}^*\hat{M} = \hat{M}\hat{M}^*$, назива се „нормалном“, а њене својствене вредности леже на јединичној кружници. Према томе, унитарна матрица је нормална. Због $\hat{U}^*\hat{U} = \hat{I}$, редови (колоне) ове матрице чине ортогоналну базу простора вектора са комплексним коефицијентима. Додатно, зато што је иначе детерминанта производа матрица једнака производу њихових детерминанти, унитарна матрица је изометрија, она не мења интензитет вектора деловањем.

Посебан случај су унитарне матрице које коњуговање не мења. Код обичних бројева, матрица реда $n = 1$, ако је $z = z^*$ онда је број z реалан. Код матрица реда $n = 2$ и већег, коефицијенти могу бити комплексни, иако је $\hat{U} = \hat{U}^*$, као што ћемо видети у следећем примеру код Паулијевих матрица. При томе приметимо да су све унитарне матрице неки корене јединичне матрице, а то су нарочито матрице једнаке својим коњугованим. Унитарне матрице се у том смислу могу сматрати генерализацијама Паулијевих.

Пример 3. Паулијеве матрице су квадратне другог реда које решавају једначину $\sigma^2 = \hat{I}$. То су:

$$\hat{\sigma}_x = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix},\ \hat{\sigma}_y = \begin{bmatrix}0 & -i\\i & 0\end{bmatrix},\ \hat{\sigma}_z = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$$

што је лако проверити. Оне се много користе у квантној физици. ❑

Пример 4. Паулијеве матрице су решење матричне једначине $\hat{\sigma}^2 = \hat{I}$, а кватерниони

решење једначине $\hat{q}^2 = -\hat{I}$. Ако прву решава матрица $\hat{\sigma}$, другу решава $\hat{q} = i\hat{\sigma}$. Према томе,

уместо матрица из претходног примера имамо четири матрице:

$$\hat{q}_x = \begin{bmatrix}0 & i\\i & 0\end{bmatrix},\ \hat{q}_y = \begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix},\ \hat{\sigma}_z = \begin{bmatrix}i & 0\\0 & -i\end{bmatrix}$$

које називамо „кватернионима“. ❑

Детаље о Паулијевим матрицама и кватернионима погледајте у мојој књизи „Квантна Механика“, на пример стр. 129.

II Део: Информација перцепције

Идеја „информације перцепције'' у мојој истоименој књизи (2016) је наглашавање опажања. Оно што се „дешава" је оно што интерагује, што шаље или прима физичка дејства, а да је то уједно оно што размењује (физичке) информације, односно комуницира.

Да би идеју тог „дејства" осмислили као „информацију", треба нам „објективност случајности" и мерење „количине опција". Ови појмови као да постоје у познатим теоријама, али приметићете они немају тачно она значења која нам сада требају. Потребне су нам боље поставке веза онога што називамо реалношћу, неизвесношћу, могућностима и опажајима. Све ћемо их подвести под информације.

Ако A и B могу „комуницирати", тада и само тада рећи ћемо да су A и B узајамно „реални". Када они могу непосредно комуницирати (шта год то значило), непосредно су реални. Када A може комуницирати са C, а C са B, али нема комуникације између A и B, нема непосредне комуникације и нема непосредне реалности. Ово свођење реалности на размену информација лако се преноси на еквиваленцију информације на интеракцију, важан корак у теорији информације перцепције.

Информацију посматрамо као градивно ткиво простора, времена и материје, а саму њену суштину видимо у неизвесности. Већ због тога и поменуте дефиниције реалности, неизвесност мора бити и објективна и релативна. Прво значи постојање макар неких несагледивих опција за сваки субјекат опажања реалности, а друго значи зависност „несагледивости" од датог субјекта. За схватање овог „апсурда" наводио сам пример ловца који лови клопкама и лукавошћу дивљач, која при томе није у стању предвидети (разумети) шта јој се спрема.

То је концепт којем још „само" треба математичка форма. Зашто математичка – питао ме је једном колега, алудирајући на допринос филозофији – зато што не познајемо никакво друго умовање које би могло филтрирати апсолутне истине. Метода дедукције без методе контрадикције јалова је, а ова друга, на ваљан начин, користи се једино у математици.

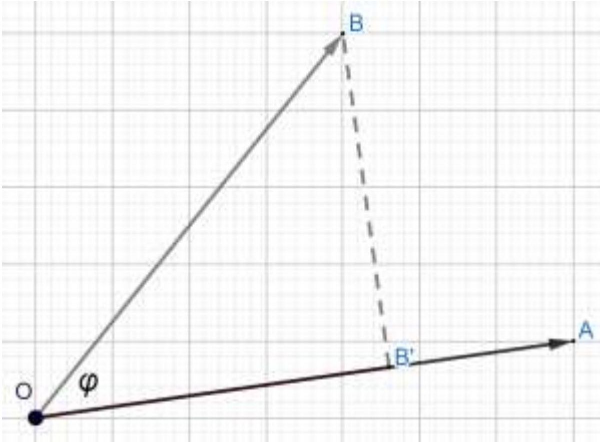
Елем, количину перцепције уоченог субјекта у датим околностима, дакле неки претпостављени обим размене информације таквог са околином, назовимо „информацијом перцепције" (S), моћ контроле опција датог субјекта назовимо „интелигенцијом" (I), а способност околине која га при томе спречава, спутава „хијерархијом" (H). Наводнике подразумевамо, јер ови појмови нису баш они на које обично мислимо, јер знамо да честица-талас квантне механике нема интелигенцију ни рецимо друштвену хијерархију у којој живи.

Даље, приметимо да способност субјекта да користи различите опције (тзв. интелигенција) расте са моћи перцепција (S), а опада са ограничењима наметнутм од стране окружења (H). Упрошћено можемо ставити $I = S : H$, односно $S = I \cdot H$. Међутим, околина може бити сложенија, а и субјект опажања, па је ову формулу потребно видети као део шире $S = I_1 \cdot H_1 + I_2 \cdot H_2 + I_3 \cdot H_3 + \cdots$, где су фактори сабирака I_k и H_k, за $k = 1, 2, 3, \ldots$, претходно описаног смисла, евентуално парцијални па и међусобно (за различите индексе k) независни.

18. Унитарни простори

Унитарне операторе линеарна алгебра дефинише на тзв. унитарним просторима, а ови су, просто речено, векторски простори снабдевени скаларним производом вектора. Више пута сам наводио те додатне аксиоме које векторски простор могу учинити „унитраним“, а сличне можете налазити свукуда. Ево неколико примера.

Пример 1. Нека је X векторски 3-дим простор усмерених дужи Декартовог правоуглог система координата $Oxyz$. Дата су два вектора $\vec{a} = a_x\vec{\imath} + a_y\vec{\jmath} + a_z\vec{k}$ и $\vec{b} = b_x\vec{\imath} + b_y\vec{\jmath} + b_z\vec{k}$, где су $\vec{\imath}, \vec{\jmath}$ и $\vec{k}$ јединични вектори координатних оса (ортови). Краће пишемо $\vec{a} = (a_x, a_y, a_z)$ и $\vec{b} = (b_x, b_y, b_z)$.



На слици лево, $\overrightarrow{OA} = \vec{a}$, $\overrightarrow{OB} = \vec{b}$, $\sphericalangle(\vec{a}, \vec{b}) = \varphi$, што је увек могуће, јер ма колико димензија имао систем координата два вектора леже у једној равни.

Окомита, тзв. ортогонална пројекција вектора $\overrightarrow{OB}$ на вектор $\overrightarrow{OA}$ је $\overrightarrow{OB'}$, интензитета

$$\left|\overrightarrow{OB'}\right| = \left|\overrightarrow{OB}\right| \cos\varphi$$

па је производ интензитета пројекције једног вектора на други и другог

$$\overrightarrow{OA} \cdot \overrightarrow{OB} = \left|\overrightarrow{OA}\right|\left|\overrightarrow{OB}\right| \cos\varphi$$

Исти резултат даје и интензитет пројекције вектора $\overrightarrow{OA}$ на $\overrightarrow{OB}$ множен са $\left|\overrightarrow{OB}\right|$.

Скаларни производ ових вектора (усмерених дужи) је производ њихових интензитета и косинуса угла између њих, $\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos\sphericalangle(\vec{a}, \vec{b})$, а рачун даље показује да се исто може написати само помоћу координата $\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + a_z b_z$. У средњој школи се учи зашто су те две вредности увек једнаке. ❑

Коментарисаћу укратко тачност ове дефиниције у односу на аксиоме унитарних простора, према дефиницији 1.2.46 моје књиге „Квантна механика“.

1. Коњугована симетрија: $(\vec{a} \cdot \vec{b})^* = \vec{b} \cdot \vec{a}$, очигледна је обзиром да су коефицијенти реални.
2. Адитивност по другом аргументу: из дистрибутивности, $\vec{a} \cdot (\vec{b}_1 + \vec{b}_2) = \vec{a} \cdot \vec{b}_1 + \vec{a} \cdot b_2$.
3. Хомогеност по другом аргументу: због $\vec{a} \cdot (\lambda\vec{b}) = \lambda\vec{a} \cdot \vec{b}$.
4. Позитивна дефинитност: $\vec{a} \cdot \vec{a} = a_x^2 + a_y^2 + a_z^2 \ge 0$, резултат је такође позитивне дефинитности квадрата реалних бројева, а из истог следи и последња,
5. аксиома о само-окомитости: $\vec{a} \cdot \vec{a} = 0 \Leftrightarrow \vec{a} = 0$.

Додатно, приметимо да скаларни производ вектора $p(\vec{a}, \vec{b}) = \vec{a} \cdot \vec{b}$ расте са смањивањем угла φ између њих, јер косинус тада расте, а увек је $p(\vec{a}, \vec{b}) \le |\vec{a}||\vec{b}|$, због $|\cos\varphi| \le 1$. А једнакост важи ако и само ако $\varphi = 0$, тј. ако су вектори истог правца. Ово значи да векторима јединичне дужине скаларни производ није већи од јединице и да је могућа њихова вероватносна интерпретација. Мањи угао између њих интерпретираћемо бољом „адаптацијом“, а раст скаларног производа

тада значиће пораст вероватноће у случају „удруживања“, дакле доследно принципима теорије информације.

Када скаларни производ примењујемо на склоност природе ка вероватнијим исходима, односно на развој ка мањој информацији, подсећам да [05.] производ $p = p(\vec{a}, \vec{b})$ увек можемо разумети као вероватноћу $p = 1/n$ неких n једнако вероватних исхода и израчунати информацију на начин Хартлија. Тада њих „замишљених“ n она износи $h = \log_b n = -\log_b p$. База b логаритма одређује јединицу мере информације, рецимо bit за $b = 2$, или nat за $b = e \approx 2{,}71828$.

Пример 2. Нека су μ_1, μ_2, μ_3 позитивни бројеви и X векторски простор орјентисаних дужи, као у претходном примеру, тада је $\langle \vec{a}, \vec{b} \rangle = \mu_1 a_x b_x + \mu_2 a_y b_y + \mu_3 a_z b_z$ врста скаларног производа вектора. ❑

Слично претходном и овде се лако проверавају аксиоме унитарних простора. Коњугована симетрија, $\langle \vec{b}, \vec{a} \rangle^* = \langle \vec{a}, \vec{b} \rangle$, очигледна је због комутације множења и реалности коефицијената. Такође адитивност, $\langle \vec{a}, \vec{b}_1 + \vec{b}_2 \rangle = \langle \vec{a}, \vec{b}_1 \rangle + \langle \vec{a}, \vec{b}_2 \rangle$, као и хомогеност, $\langle \vec{a}, \lambda\vec{b} \rangle = \lambda \langle \vec{a}, \vec{b} \rangle$, по другом аргументу. Позитивна дефинитност, $\langle \vec{a}, \vec{a} \rangle \geq 0$, следи из позитивности коефицијената μ_k и особина из претходног примера, а отуда и само-окомитост нуле, $\langle \vec{a}, \vec{a} \rangle \Leftrightarrow \vec{a} = 0$.

Приметимо да се скаларни производи лако преводе у метрике, мерења интервала, или растојања између тачака. Тада они не постају већи када су вектори боље „адаптирани“, већ напротив мањи. Да би се то десило потребна је узајамна зависност вектора спрегнутих у скаларном производу, да међу факторима сабирака „више једног значи мање другог“. Очигледно је да то може и покварити игрaчку, али да такве (не приметно на овај начи) већ постоје у физици. На пример, Хајзенбергове релације неодређености су такав случај.

Наиме, ако су компоненте првог вектора неодређености импулса Δp_x, Δp_y и Δp_z, дуж праваца чије су неодређености Δx, Δy и Δz компоненте другог вектора, онда „скаларни производ“ тих вектора $s = \Delta p_x \cdot \Delta x + \Delta p_y \cdot \Delta y + \Delta p_z \cdot \Delta z$ тежи узимању минималних вредности, јер већа неодређеност импулса таласа-честице значи мању неодређеност њеног положаја и обрнуто. Међутим, како је поједино Δp мање што је Δx веће, то не важи друга од аксиома скаларног производа (адитивност), али се помоћу овог може дефинисати $p = e^{-s}$, тако да је $s = -\ln p$, при чему се ово p понаша као вероватноћа и у складу са аксиомама скаларног производа.

Пример 3. У n-димензионалном векторском простору n-торки комплексних бројева C^n, векторе $\vec{x} = [\![\xi_k]\!]$ и $\vec{y} = [\![\eta_k]\!]$ удружимо изразом $\langle \vec{x}, \vec{y} \rangle = \xi_1 \eta_1^* + \xi_2 \eta_2^* + \dots + \xi_n \eta_n^*$ и он постаје исправно дефинисан скаларни производ. Приметимо да без коњуговања другог фактора (у сваком од сабирака) дефиниција не би ваљала. ❑

Пример 4. Векторски простор непрекидних функција $x, y \in C(a, b)$ на коначном интервалу $[a, b]$ има скаларни производ интеграл $\langle x, y \rangle = \int_a^b x(t) y^*(t)\ dt$. ❑

Пример 4. је пренешен 3. пример на збирове (интеграл) инфинитезималних величина.

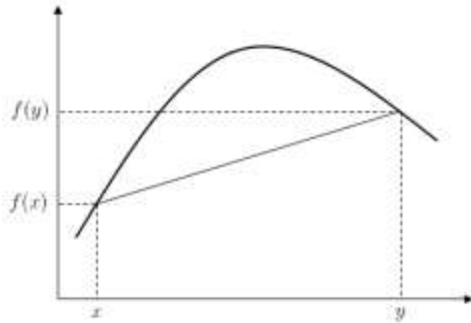
19. Метрички простори

Близак појам векторском простору је метрички простор. Њега чини аморфан скуп X који сада називамо тачкама (претходно векторима). Тачке су n-торке бројева, односно координата, као на пример $\vec{x} = (\xi_1, \xi_2, \dots, \xi_n)$ и $\vec{y} = (\eta_1, \eta_2, \dots, \eta_n)$, чијем је сваком уређеном пару $(\vec{x}, \vec{y})$ придружен број $d(\vec{x}, \vec{y})$, тзв. дистанца, или удаљеност међу њима, за који важе следеће аксиоме:

1. Идентичност: $d(\vec{x}, \vec{y}) = 0 \Leftrightarrow \vec{x} = \vec{y}$;
2. Позитивност: $0 \le d(\vec{x}, \vec{y}) < +\infty$;
3. Симетрија: $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$;
4. Неједнакост троугла: $(\forall \vec{z} \in X)\ d(\vec{x}, \vec{y}) \le d(\vec{x}, \vec{z}) + d(\vec{z}, \vec{y})$.

Редослед аксиома је дефиниције 1.3.1 из моје књиге „Квантна Механика“, стр. 189. Тамо је речено пуно тога што бих овде могао преписати, а само нешто од тога ћу споменути. Пре свега, редом су то доказане неједнакости: Јангова (лема 1.3.2), Хелдерова (теорема 1.3.3) и Минковског (теорема 1.3.4). Затим још неколико за квантну механику важних теорема функционалне анализе, као и примери метричких простора.

За конкавне функције важи да се граф налази изнад дужи која спаја њене две тачке, као на слици



$$(1-t)f(x) + tf(y) < f\big((1-t)x + ty\big).$$

Логаритамска функција је конкавна, па је

$$(1-t)\log a^{\frac{1}{1-t}} + tb^{1/t} < \log\left((1-t)a^{\frac{1}{1-t}} + tb^{1/t}\right)$$

за све $a, b \ge 0$. На левој страни је $\log(ab)$, а отуда

$$ab \le \frac{a^p}{p} + \frac{b^q}{q},$$

где $p > 1$, $q > 1$ и $\frac{1}{p} + \frac{1}{q} = 1$. Једнакост важи ако и само ако $a^p = b^q$. То је Јангова неједнакост.

Стављајмо у Јангову неједнакост $a = \left|\alpha_k / \left(\sum_j |\alpha_j|^p\right)^{1/p}\right|$ и $b = \left|\beta_k / \left(\sum_j |\beta_j|^q\right)^{1/q}\right|$ редом за индексе $k = 1,2,\dots,n$ и саберимо

$$\sum_{k=1}^{n} |\alpha_k \beta_k| \le \left(\sum_{k=1}^{n} |\alpha_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |\beta_k|^q\right)^{1/q}$$

где је $p > 1$ и $\frac{1}{p} + \frac{1}{q} = 1$. То је Хелдерова неједнакост. Два пута примењујући Хелдерову, добијамо неједнакост Минковског:

$$\sum_{k=1}^{n} |\alpha_k + \beta_k|^p \le \sum_{k=1}^{n} |\alpha_k||\alpha_k + \beta_k|^{p-1} + \sum_{k=1}^{n} |\beta_k||\alpha_k + \beta_k|^{p-1} \le$$

$$\le \left(\sum_{k=1}^{n} |\alpha_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |\alpha_k + \beta_k|^{qp-q}\right)^{\frac{1}{q}} + \left(\sum_{k=1}^{n} |\beta_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |\alpha_k + \beta_k|^{qp-q}\right)^{\frac{1}{q}}$$

$$\leq \left[\left(\sum_{k=1}^{n}|\alpha_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n}|\beta_k|^p\right)^{\frac{1}{p}}\right]\left(\sum_{k=1}^{n}|\alpha_k+\beta_k|^{qp-q}\right)^{\frac{1}{q}}$$

што делећи другим фактором даје

$$\left(\sum_{k=1}^{n}|\alpha_k+\beta_k|^p\right)^{1/p} \leq \left(\sum_{k=1}^{n}|\alpha_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n}|\beta_k|^p\right)^{\frac{1}{p}}$$

а то је неједнакост Минковског. Са неједнакошћу Минковског, углавном, доказујемо неједнакост троугла, а тиме и метрику, односно да су одређени скупови метрички простори.

Пример 1. Простор C_r^n има метрику $d(\vec{x},\vec{y}) = \left(\sum_{k=1}^{n}|\xi_k - \eta_k|^r\right)^{1/r}$, где $1 \leq r < \infty$. ❑

Ако су низови реалних бројева то је простор R_r^n, у случају $r = 2$ доњи индекс се не мора писати.

Пример 2. Када $r \to \infty$ ови простори прелазе у C_∞^n метрике $d(\vec{x},\vec{y}) = \max\limits_{1\leq k\leq n}|\xi_k - \eta_k|$. ❑

Пример 3. Када су тачке метричког простора ограничени низови бројева, а растојање уведено са $d(\vec{x},\vec{y}) = \sup\limits_{1\leq k\leq\infty}|\xi_k - \eta_k|$, простор је ознаке m. ❑

Пример 4. Када $n \to \infty$, тачке $x = [\![\xi_k]\!]$ простора l_r, за $1 \leq r < \infty$, су бесконачни низови бројева такви да ред $\sum_{k=1}^{\infty}|\xi_k|^r$ конвергира. Растојање је $d(x,y) = \left(\sum_{k=1}^{\infty}|\xi_k - \eta_k|^r\right)^{1/r}$. Када је $r = 1$ пишемо l. Јасно је да простор C_r^n прелази у l_r када $n \to \infty$. ❑

О начинима дефинисања инфрмације перцепције метричких простора писао сам популарно у блогу недавно (Functional), али том приликом нисам помињао начин „стереометрије статистике“ који сам делом описао већ 1916. године у књизи „Информација перцепције“. Такав би текст био претежак за блог, ко би га тамо разумео, ако је и наведени разумљив.

Наиме, идеја је да се тачке које су, попут $x = (\xi_1, \xi_2, \xi_3, \dots)$, неке n-торке или бесконачни низови реалних или комплексних бројева, препознају као случајне варијабле. Метрика таквог простора представља врсту дисперзије. Удаљеност $d(x,y)$ постаје мера одступања, расипања случајних вредности, а метрички простори добрим делом темама теорије информације, вероватноће и статистике. Кроз следеће примере ћемо разумети ширину ове методе.

Пример 5. Замислимо да имамо низ од $n = 1, 2, 3, \dots$ било каквих унапред познатих случајних догађаја. То могу бити падање „главе“ након бацања новчића, падање „шестице“ након бачене коцке, извлачење „пик аса“ из шпила карата и слично. Нека је сваки од тих догађаја један из скупа Ω = $\{\omega_1, \omega_2, \dots, \omega_n\}$, са вероватноћом $p_k = P(\omega_k)$.

Ове вероватноће су реални бројеви између 0 и 1, али им збир не мора бити један. Оне не чине тзв. расподелу верованоћа и биће само пример за разумевање „метрике вероватноћа“.

Један циклус нека је по један покушај повољног исхода сваког од припремљених догађаја. Након $M = 1, 2, 3, \dots$ циклуса имаћемо број m_k успешних исхода сваког k-тог ($k = 1,2, \dots, n$) догађаја. Тај

подељен са M даје број $\xi_k = m_k/M$, тзв. релативну учесталост датог исхода. Након M циклуса имамо догађаје (ω_k), вероватноће (p_k) и релативне учесталости (ξ_k) у редовима табеле:

$$\hat{X}_M = \begin{pmatrix} \omega_1 & \omega_2 & \dots & \omega_n \\ p_1 & p_2 & \dots & p_n \\ \xi_1 & \xi_2 & \dots & \xi_n \end{pmatrix}.$$

Овако добијамо низове случајних варијабли $\vec{x} = (\xi_1, \xi_2, \dots, \xi_n)$ и $\vec{y} = (\eta_1, \eta_2, \dots, \eta_n)$, које су релативне учесталости. Они су различити у различитим корацима M. Штавише, ако различитих дана крећемо испочетка са истим циклусима, након једнаких бројева корака можемо добијати различите низове $\vec{x}$ и $\vec{y}$. Растојања $d(\vec{x}, \vec{y})$ врсте су дисперзија разних метрика. ❑

Са већим бројем циклуса, из $M \to \infty$ следи $d(\vec{x}, \vec{y}) \to 0$, без обзира на избор метрике, кориштене на начин претходног примера. Ова конвергенција последица је закона великих бројева теорије вероватноће, а њена последица да су пресликавања попут $\vec{x}(M) \to \vec{x}(M+1)$ контракције простора (Банахов став). Непокретна тачка таквог пресликавања, заснованог на примеру 5, је низ вероватноћа $\vec{p} = (p_1, p_2, \dots, p_n)$.

Приметимо да релативне учесталости (број повољних исхода подељен бројем покушаја), са све већим бројем покушаја, теже ка вероватноћама исхода аналогно „адаптацији" каналом [11.], те да и овај начин препознаје начело минимализма информације. Ово наглашавам због ранијих мојих објашњења (Infinity II) да је свемир у сталној креацији, а не само у „рекреацији" једном заувек створених стања. Тај мој став био је својевремено сасвим нов, неочекиван и ваљда зато одбојан (колегама), да би га већ данас проналазио којекуда, додуше углавном извевeден, тумачен или употребљен на начине са којим се ни ја не слажем.

Зато подвлачим да овде требамо обратити пажњу на додатне могућности адаптације природних процеса и стања. Поред конвергенције поруке проласком кроз низ (истих) канала ка својственим величинама датог канала, или смањивања неизвесности преласком из микро у макро свет, у духу је теорије информације перцепције посматрати и адаптације самих канала преноса порука. Такође и мењање иначе окружења (информација је ткиво свега), од адаптације природних појава око нас на оно у шта су оне претходно еволуирале, до прилагођавања ширег свемира. Посебно би требало размотрити и физичко време кроз последице ових усклађивања, рецимо као процес понављања који законима великих бројева смањује (локалну) текућу информацију.

Како је могуће имати креацију свемира обзиром на закон одржања, рећи ћу, тако што је важнија количина него суштина (Farming).

20. Банахов простор

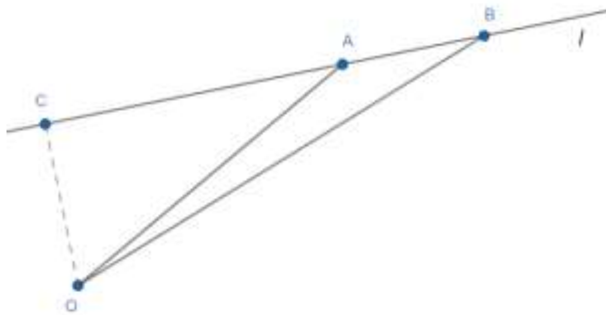
Комплетан нормиран векторски простор, у математици, тачније у функционалној анализи, назива се Банахов простор. Реална функција са следећим особинама, аксиомама ($\forall\vec{x},\vec{y} \in X$ и $\forall\lambda$ скалар), назива се норма вектора:

1. $\|\vec{x}\| \geq 0$ – ненегативност;
2. $\|\lambda\vec{x}\| = \lambda\|\vec{x}\|$ – хомогеност;
3. $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ – неједнакост троугла.

У унитарном простору, норму вектора дефинише корен скаларног производа вектора са самим собом, $\|\vec{x}\| = \sqrt{\vec{x}\cdot\vec{x}}$. У метричком простору, норма је удаљеност од нултог низа, $\|\vec{x}\| = d(\vec{x},0)$. Лако је тада проверити исправност горњих особина норме. Обрнуто, када имамо нормиран простор, метрику дефинише једнакост $d(\vec{x},\vec{y}) = \|\vec{x} - \vec{y}\|$.

У метричком простору (X,d) за низ $\vec{x}_1$, $\vec{x}_2$, $\vec{x}_3$, … кажемо да је Кошијев ако за сваки позитиван реалан број $\varepsilon > 0$ постоји цели број $N > 0$ тако да за све целе бројеве $m,n > N$ важи неједнакост $d(\vec{x}_m,\vec{x}_n) < \varepsilon$. Метрички простор X назива се комплетан, потпун, или Кошијев простор, ако сваки Кошијев низ тачака у X има лимес, граничну вредност, која је такође у X.

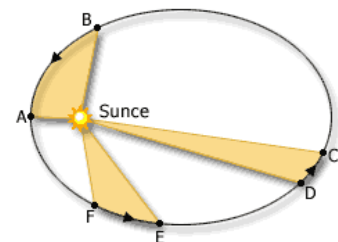
Пример 1. Посматрајмо линију l по којој се „крећу“ две тачке $A,B \in l$ и тачку O ван те линије, на слици лево. Нормала из O на линију l пада у тачку $C \in l$. Када су l и тачка O фиксирани, не мењају узајамни положај, онда троугао ABO има увек исту површину акко[2] $d(A,B)$ = const.



Подразумевамо да су тачке слике елементи неког метричког простора (X,d) и да су дате дужине $\overline{AB} = d(A,B)$, $\overline{OA} = \|A\|$, $\overline{OB} = \|B\|$, $\overline{OC} = \|C\|$. Површина ΔABO је полу производ основице $\overline{AB}$ и њене висине $\overline{OC}$. ❑

Тачке $A,B \in l$ горњег примера можемо сматрати једном тачком која се из прве позиције (A) креће по линији (l) и након неког времена ($\Delta t = t_2 - t_1$) стиже у другу позицију (B), тако да за једнака времена прелази једнака растојања ($\overline{AB}$). Константну површина троугла ABO, при томе, даље схватимо одсуством централне силе у тачки O, односно не поседовањем набоја покретне тачке.

Пример 2. Други Кеплеров закон, на слици десно, каже да потег од сунца до планете, која се креће под дејством гравитације сунца, у истим временима пребрише једнаке површине. Овај се резултат може поопштавати, али не тако да важи и обрнуто (Прилози, Централно кретање).



Наиме, не само гравитациона већ свака централна сила има исто својство, да набој покреће по некој линији при чему ће потег од ње до набоја у једнаким временима пребрисати једнаке површине. ❑

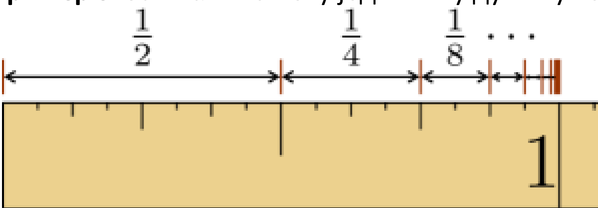
[2] акко – ако и само ако.

Са избором дефиниције метрике и норме, из модела које већ имамо у алгебри и функционалној анализи, пример 1 и пример 2 могу се разумети као посебни су случајеви апсолутне геометрије (Еуклидове, Лобачевског, Риманове), затим Ајнштајнове опште теорије релативности. Овде имам намеру испитати ширење те идеје у теорији информације, у разматрању „пребрисане површине“ као информације и њене „силе“, односно закона одржања и принципијелног минимализма.

У представљеном циљу остаје отворено питање комплетности простора и времена које ћемо прво расправити. Диференцијална геометрија, на којој је у Прилозима I-III заснивана идеја примера 1-2, дакле у наставку актуелну, не може без непрекидности. Са друге стране (Packages) информација је еквивалентна квантима дејства, најмањим пакетима који заузимају простор-време, али кретањем тамо-овамо не успевају бити „мањи од унапред датог $\varepsilon > 0$“. Међутим, може се показати да је ова наизглед контрадикција заправо привидна, објашњива самом дефиницијом реалности на начин теорије информације.

Теорија информације перцепције реалност своди на комуникацију, уопште на спрегу субјекта који опажа и опажаног објекта, чије су улоге замењиве. Са тог становишта можемо рећи да увек около постоји неки посматрач који дату величину може видети довољно малом, мањом од унапред дате мале величине (дужине, трајања, енергије, импулса и сличног). А управо то је игра епсилон-делта инфинитезималног рачуна.

Пример 3. Замислимо неку јединичну дужину коју треба прелазити напредујући у порцијама по пола од преосталог, као на слици лево. Пређе се прво половина, па четвртина, па осмина, па пола од преосталог (шеснаестина) и тако без краја, бивајући све ближе крају а никада га не достижући. Тако некако почињу све приче о инфинитезималном рачуну у уџбеницима.



Парцијални збир, суму $s_n = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^n}$, посматрамо као низ $s_1, s_2, s_3, \dots, s_n, \dots$ који је за све веће $n = 1, 2, 3, \dots$ све ближи вредности 1.

Када постоји број s такав да за свако $\varepsilon > 0$ постоји довољно велики природни број n_0 који ће за сваки индекс $n > n_0$ учинити тачном неједнакост $|s_n - s| < \varepsilon$, онда кажемо да низ s_n конвергира ка s и пишемо

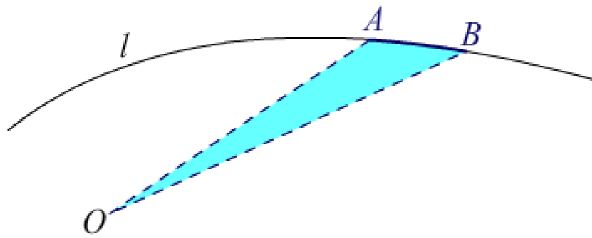
$$\lim_{n\to\infty} s_n = s$$

односно $s_n \to s$ када $n \to \infty$. У датом случају $s = 1$. ❑

Приметићу, још једном, оснивачи диференцијалног рачуна као да су намерно тражили баш такву дефиницију граничне вредности (лимеса), на коју бисмо могли надовезати теорију информације, упркос квантовања (атомизирања) њених најмањих пакета. Тај привид, да су они „натрапали“ на управо овако нама сада згодну дефиницију, заправо је последица нечега много дубљег. Истине математике увек су тако згодне, неконтрадикторне са собом, као и са теоријом којој припадају, или са било којим делом познате или непознате друге тачне теорије, да нам се сваки пут могу учинити као намештаљка.

21. Проблем површине

Набој се под дејством константне централне силе O креће $A \to B$ дуж линије l. Поново нам треба претходна ситуација [20. Пример 1-2], која је рекох детаљ много раније анализе чије делове сам сачувао у скрипти „Прилози“. Циљ су ми биле особине централне силе изведене путањама набоја и обрнуто, путање из сила. Резултат су конике, криве другог реда, за силу која опада са квадратом удаљености, чији посебан случај је Кеплеров други закон.



Једнаке површине које пребрише потег од сунца до планете (од OA до OB) у једнаким временима одговарају сталном дејству гравитационе силе (O) на дату планету ($A \to B$). Увек једнака површина троугла (OAB) фиксираног темена (O), чија основица (AB) клизи дуж тада праве линије (l), указује на слободно кретање без дејства силе. Такве и сличне ситуације говоре о константном физичком дејству (енергија × време) константне централне силе.

Како је дејство еквивалент информацији, то у једнаким трајањима ове једнаке површине говоре о еквивалентним комуникацијама константне силе и набоја коју она покреће. У одговарајућим тим условима површина би могла одговарати информацији, бити веза као између облика путање (l) и силе (O), односно обрнуто.

Важност површине за информацију видимо у случају чула и перцепције[3], или код пропадања тела у јако гравитационо поље црне рупе. Са становишта релативног посматрача, негде даље, дужине у правцу силе (радијалне) тог тела скраћују се, а време телу успорава. Оно се размазује око извора силе и постаје површина сфере изнад хоризонта догађаја (границе црне рупе), а сва информација тела постаје површина сфере у неком безвременом стању. Такође, информација као површина, а тачније речено као гранична област, приметна је и у доказу паралелних реалности (Dimensions), на начин „теорије информације“ (моје).

То је „метода површине“ као и раније овде поменута [4.] ради повезивања тригонометрија, које ће се показати корисним касније у теорији информације перцепције. Али математичке аналогије на једном месту откривене виђаћемо свукуда. Таква је природа апстракција које говоре о истини, да су неизбежне и скоро посвуда присутне, али као да их природа не воли (The Truth).

Тако поред физичких сила (иначе дискутабилног концепта[4]) неку врсту привлачности има и начело вероватноће чинећи чешћим вероватније исходе, па онда и одговарајући информације, да стања теже ка мање информативнима, или принцип најмањег дејства физике. То су опште спонтаности са многобројним последицама на које ће се пренети аналогије „силе“. Међу њима је регресиона анализа, али и решења „задатака површине“ у повезивању расподела случајних промењивих са метрикама (за сада непознат начин), као и тек поменута [19. пример 5] конвергенција релативних вероватноћа са повећањем броја понављања опита.

Вратимо се поменутом примеру [19. 5], али сада без нормирања. Низ од $n = 1, 2, 3, \ldots$ било каквих унапред познатих случајних догађаја тестирамо сваки по M пута и бележимо бројеве сада рецимо

[3] Р. Вуковић: „Природа Података“, Нови Глас Бања Лука, 1999, Чула стр. 21-28.
[4] Р. Вуковић: „Приче о информацији“, 2.19 Класична сила

ξ_k успешних исхода сваког k-тог ($k = 1,2, \dots, n$) догађаја. Са повећањем бројева M циклуса бројеви $\xi_k \le M$ расту (не делимо их бројем покушаја M) и не представљају релативне вероватноће као у оригиналном примеру.

Низ $\vec{x} = (\xi_1, \xi_2, \dots, \xi_n)$ је вектор у неком n-димензионалном простору могућности. Када поновимо циклусе испочетка и дођемо опет до новог M-тог покушаја, добићемо нови низ $\vec{y} = (\eta_1, \eta_2, \dots, \eta_n)$ са вероватно другим координатама, у истом n-дим простору могућности и приближно исте норме, $\|\vec{x}\| \approx \|\vec{y}\|$. Са порастом броја покушаја, ове норме расту, а поједине координате дељене текућим бројем постају релативне вероватноће које се групишу око вероватноће, $\xi_k/M \to p_k$ када $M \to \infty$. Тада $\vec{x}, \vec{y} \to \vec{p} = (p_1, p_2, \dots, p_n)$ и $d(\vec{x}, \vec{y}) \to 0$.

Ма колико да је димензионалан простор могућности, колико год координата да имају вектори $\vec{x}$ и $\vec{y}$, они разапињу један паралелограм, суседних страница $\|\vec{x}\|$ и $\|\vec{y}\|$ са неким углом φ између. Из тога следи њихов скаларни производ

$$\vec{x} \cdot \vec{y} = \|\vec{x}\|\|\vec{y}\| \cos\varphi = \xi_1\eta_1 + \xi_2\eta_2 + \dots + \xi_n\eta_n = \sum_{k=1}^{n} \xi_k\eta_k$$

а из овога псеудо-скаларни производ $|\vec{x} \times \vec{y}| = \|\vec{x}\|\|\vec{y}\| \sin\varphi$ који би представљао површину.

У аналогији са горњим (кретањем под дејством константне централне силе), понављајући циклусе и бирајући M = const, очекујемо изразити ове површине (троуглова $O\vec{x}\vec{y}$). Води нас идеја да закон одржања енергије има неку корисну аналогију са одржањем вероватноћа. Да важи и такав закон одржања „мере неизвесности“. Са повећањем броја понављања опита ($M \to \infty$) ове координате се стабилизују ($\xi_k/M \to p_k$), а сви вектора $\vec{x}$ правци конвергира ка правцу вектора $\vec{p}$. Странице троугла ($\|\vec{x}\|$ и $\|\vec{y}\|$) постају дуже, али угао између њих (φ) се смањује, све мањи је синус тог угла ($\sin\varphi$) који је трећи фактор површине (заједно са две странице), што сужава могућности мењања површине.

Како и са истим бројевима (M) циклуса можемо добијати значајно различите површине речених троуглова ($O\vec{x}\vec{y}$), бирајући различите метрике $d(\vec{x}, \vec{y})$ и према томе различите норме $\|\vec{x}\| = d(\vec{x}, 0)$ и $\|\vec{y}\| = d(\vec{y}, 0)$, тако можемо и обрнуто, различите површине оваквих узастопних троуглова исте метрике сводити на сличне површине различитих метрика, донекле. То је пример аналогије коју горе најављујем, назовимо је „проблем површине“.

22. Неједнакост метрика и аналогија

Вратимо се [19. Пример 1] на простор C_r^n метрике $d(\vec{x},\vec{y}) = \left(\sum_{k=1}^{n}\left|\xi_k - \eta_k\right|^r\right)^{1/r}$. Ови вектори, или низови $\vec{x} = (\xi_1, \xi_2, \dots, \xi_n)$ и $\vec{y} = (\eta_1, \eta_2, \dots, \eta_n)$ су реални или комплексни бројеви, а координате су неког n димензионалног простора могућности, рецимо [19. Пример 5]. Приметимо да без губитка општости, увек можемо издвојити позитивне сабирке, $a_k = \left|\xi_k - \eta_k\right| > 0$, изостављајући нулте.

Теорема 1. Функција $f(x) = (a_1^x + a_2^x + \dots + a_n^x)^{1/x}$ са позитивним константама a_k и $x > 0$ већа је од највећег сабирка и опадајућа је.

Доказ: Нека је $a_\beta = \max_{1\le k\le n}\{a_k\}$ поменути „највећи сабирак“. Очигледно је

$$f(x) = a_\beta\left(1 + \sum_{k\neq\beta}\frac{a_k^x}{a_\beta}\right)^{\frac{1}{x}} > a_\beta$$

јер су сви сабирци позитивни бројеви. Тиме је први део теореме доказан.

Нека је $a_\alpha = \min_{1\le k\le n}\{a_k\}$ и пишемо $f(x) = a_\alpha(b_1^x + b_2^x + \dots + b_n^x)^{\frac{1}{x}} = a_\alpha\, g(x)$, где је $b_k = \frac{a_k}{a_\alpha} \ge 1$ за свако $k = 1, 2, \dots, n$. Функције f и g су непрекидне, јер су композиције непрекидних функција, а извод ове друге је негативан, што значи да је то опадајућа функција:

$$g'(x) = -\frac{1}{x^2}[\ln(b_1^x + b_2^x + \dots + b_n^x)](b_1^x \ln b_1 + b_2^x \ln b_2 + \dots + b_n^x \ln b_n) < 0$$

Зато је негативан и извод дате $f'(x) = a_\alpha g'(x) < 0$, па је и она опадајућа. ∎

Непосредна последица теореме је да је „удаљеност“ ових вектора утолико мања што метрика користи већи експонент r. У крајњем случају, када $r \to \infty$ тада $d(\vec{x},\vec{y}) \to M$, а то је метрика C_∞^n, тј. [19. 2]. Када $n \to \infty$ имамо трећи пример метрике [19. 3] где овај максимум мења супремум.

Поновићу још једном, ово је истраживање неизвесности као одбоје „силе“. Покушај упоређивања „сила могућности“ са Кеплеровим другим законом и на веома апстрактан начини повезивање те идеје са централним силама уопште, са физикалним примерима веза „облика путање“ и „облика силе“, све путем употребе (константне) површине коју вектор набоја, крећући се под дејством те силе, пребрише у једнаким интервалима времена.

Аналогно правимо везу низа случајних догађаја, на основу [19. пример 5], где члановима низа опита додајемо реализоване исходе, циклус по циклус до M и више, формирајући векторе $\vec{x}$ и $\vec{y}$ (све већих координата), који конвергирају правцу вероватноћа тих догађаја $\vec{p} = (p_1, p_2, \dots, p_n)$. Из закона великих бројева, јасно је да је угао између вектора $\vec{x}$ и $\vec{y}$ све мањи, ако је разлика у броју њихових циклса константна док су бројеви тих циклуса све већи. Идеја је да се путем сразмере површина троуглова (или паралелограма) које разапињу ови вектори успостави веза између путање (врхова) тих вектора и метрика.

Аналогија задире у апстраховање физичког појма силе у одбојну „силу“ неизвесности, чему је подршка принцип мање информације, односно веће вероватноће. Природа спонтано, чешће

реализује вероватније исходе, или што је исто, радије стања развија у мање информативна. Та њена „тенденција“ је наводна сила о којој говорим. Метрике овде расправљане нису једине.

Када број циклуса M расте, можемо управљати површином паралелограма који разапињу вектори $\vec{x}$ и $\vec{y}$ променом метрике, али и мењањем корака између циклуса. Вероватноће $\vec{p}$ нису, односно не морају чинити расподелу, али $\zeta_k = Mp_k$ јесте математичко очекивање варијабле ξ_k, а интензитет таквих $\|\vec{z}\| = d(\vec{z}, 0) = \sum_{k=1}^{n}|\zeta_k| = \sum_{k=1}^{n}|Mp_k| = M\sum_{k=1}^{n}|p_k| = M\|\vec{p}\| = M$ ако и само ако $\|\vec{p}\| = 1$, односно ако ове вероватноће чине расподелу, у датој метрици.

Пример 1. Нека је $\vec{p} = (p_1, \dots, p_n)$ расподела вероватноћа (случајних догађаја $\omega_1, \dots, \omega_n \in \Omega$), а метрика стандардна, $r = 2$. Тада је, након M понављања опита, очекивање варијабле:

$$E(\vec{x}) = [\![E(\xi_k)]\!] = [\![Mp_k]\!] = M[\![p_k]\!] = M\vec{p}$$

а означимо га и вектором $\vec{z} = M\vec{p}$. Вектор математичког очекивања $\vec{z}$ има стално исти правац $\vec{p}$ и све већи интензитет $\|\vec{z}\| = M$ са повећањем броја корака M, броја опита. Скаларни производ:

$$\vec{x} \cdot \vec{z} = \sum_{k=1}^{n} \xi_k \zeta_k = M \sum_{k=1}^{n} \xi_k p_k = M\vec{x} \cdot \vec{p} = M\|\vec{x}\|\|p\| = M\|\vec{x}\| \cos\varphi$$

где је угао $\varphi = \sphericalangle(\vec{x}, \vec{p})$, а по претпоставци $\|\vec{p}\| = 1$. Према томе, површина паралелограма разапетог векторима $\vec{x}$ и $\vec{z}$ је:

$$\Pi = \|\vec{x}\|\|\vec{z}\| \sin\varphi = M\|\vec{x}\|\sqrt{1 - \cos^2\varphi} = M\sqrt{1 - (\vec{u} \cdot \vec{p})^2}$$

где је $\vec{u} = \frac{\vec{x}}{\|\vec{x}\|}$ вектор $\vec{x}$ нормиран на јединицу. ❑

Приметимо да $\vec{u}$ представља релативну вероватноћу, фреквенцију исхода, те да се тај вектор, са повећањем броја опита M, приближава вектору вероватноће $\vec{p}$ по правцу (и интензитету), која у овом примеру представља неку расподелу. Конвергенција $\sin\varphi = \sqrt{1 - (\vec{u} \cdot \vec{p})^2} \to 0$ када $M \to \infty$, у складу је са законима великих бројева. Наиме, из особина математичког очекивања следи $E\left[M\sqrt{1 - (\vec{u} \cdot \vec{p})^2}\right] = M\sqrt{1 - (\vec{p} \cdot \vec{p})^2} = 0$, јер је $\vec{p} \cdot \vec{p} = 1$. Погледајмо сада и неколико примера заснованих на претходном.

23. Примери регресије I

За вероватноће расподеле $\vec{p} = (p_1, p_2, p_3,\ p_4)$ бираћемо четири позитивна броја јединичног збира и генерисати M случајних бројева (компјутерским програмом). Уз свако M, формирани су збирови исхода $\vec{x} = (\xi_1, \xi_2, \xi_3, \xi_4)$ поред вектора $\vec{z} = (Mp_1, Mp_2, Mp_3, Mp_4)$ и површина Π коју два вектора разапињу. У четири ($K = 4$) равномерна корака ($\Delta M =$ const.) добијамо четири пара случајних бројева (x, y), од којих је први број генерисаних корака (за добијање $\vec{x}$), а други је на основу првог израчуната површина, $y = \Pi(\vec{x})$.

Површина коју разапињу два дата вектора је $\Pi = \|\vec{x}\|\|\vec{z}\| \sin\varphi$, синус угла φ између вектора $\vec{x}$ и $\vec{z}$ израчунавамо помоћу косинуса, $\sin\varphi = \sqrt{1 - \cos^2\varphi}$, а косинус је количник њиховог скаларног производа подељен њиховим дужинама, интензитетима, $\cos\varphi = \vec{x} \cdot \vec{z} / \|\vec{x}\|\|\vec{z}\|$.

Понављајући програм за исте M добијају се различити вектори $\vec{x}$, али исти $\vec{z}$ све до промене низа вероватноћа расподеле $\vec{p} = (p_1, p_2, p_3, p_4)$. Узгред, табеле сам израчунавао компјутером, правећи разне програме (Python) по потреби које сматрам исувише једноставним за памћење, навођење овде. Како простим погледом није увек лако рећи мења ли се површина Π (од M, односно $\vec{p}$) и како, то даље проверавам линеарном регресијом.

Пример 1. За вероватноће поставимо $\vec{p} = (0.4, 0.3, 0.2, 0.1)$ и генеришимо случајне бројеве почев од $M = 100$, кораком $\Delta M = 10$. Уз свако M, добијамо збирове исхода $\vec{x}$, формирамо вектор $\vec{z}$, а та два разапињу површину паралелограма коју посебно израчунавамо:

- $M = 100$: $\vec{x} = (41, 30, 18, 11)$, $\vec{z} = (40, 30, 20, 10)$, $\Pi = 133.79$
- $M = 110$: $\vec{x} = (41, 31, 25, 13)$, $\vec{z} = (44, 33, 22, 11)$, $\Pi = 286.84$
- $M = 120$: $\vec{x} = (49, 33, 27, 11)$, $\vec{z} = (48, 36, 24, 12)$, $\Pi = 293.94$
- $M = 130$: $\vec{x} = (52, 40, 29, 9)$, $\vec{z} = (52, 39, 26, 13)$, $\Pi = 357.20$

За израчунавање линеарне регресије, зависности $y = a + bx$, површине (ознаке y) од компоненти случајног вектора (x), прегледно представљам четири равномерна корака ($K = 4$).

X	***Y***	***X***^{2}	***XY***
100	133.79	10 000	13 379
110	286.84	12 100	31 552
120	293.94	14 400	35 273
130	357.20	16 900	46 436
Σx = 460	Σy = 1 071.77	Σx^2 = 53 400	Σxy = 126 640

$$a = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{K(\Sigma x^2) - (\Sigma x)^2} = \frac{1\,072 \cdot 53\,400 - 460 \cdot 126\,640}{4 \cdot 53\,400 - (460)^2} = \frac{-1\,009\,600}{2\,000} = -504.8$$

$$b = \frac{K(\Sigma xy) - (\Sigma x)(\Sigma y)}{K(\Sigma x^2) - (\Sigma x)^2} = \frac{4 \cdot 126\,640 - 460 \cdot 1\,072}{2000} = \frac{13\,440}{2000} = 6.72$$

$$y = -505 + 7 \cdot x$$

На основу коефицијената линеарне регресије, приближно $a = -505$ и $b = 7$ добијених методом најмањих квадрата, друга колона горње табеле добија процену наведену у следећој.

X:	100	110	120	130
Y:	167	234	302	369

Ново генерисање бројева претходног примера даваће другачије коефицијенте, како генерисаних случајних површи $\Pi = f(\vec{x})$, тако и линије регресије ($y = a + bx$). ❑

Кораком једног, уместо десет опита, права линија регресије ($y = -505 + 7 \cdot x$) напредује десет пута спорије, овај коефицијент $b = 7$ био би тада десет пута мањи. Што се тиче коефицијента a, који је овде негативан, тако да за $x = 72$ и мањи број корака процењена површина y постаје нула и негативан број, што нема смисла, каже нам да је ова регресија само приближно права линија. Она је заправо крива и савија се према горе.

Пример 2. Са истим почетним условима $\vec{p} = (0.4, 0.3, 0.2, 0.1)$, од $M = 100$ и кораком $\Delta M = 10$, понављам претходни пример, а на случајан начин генерисан је нови вектор $\vec{x}$ са последицама:

- $M = 100$: $\vec{x} = (49, 33, 14, 4)$, $\vec{z} = (40, 30, 20, 10)$, $\Pi = 642.73$
- $M = 110$: $\vec{x} = (56, 29, 23, 12)$, $\vec{z} = (44, 33, 22, 11)$, $\Pi = 282.38$
- $M = 120$: $\vec{x} = (44, 43, 21, 12)$, $\vec{z} = (48, 36, 24, 12)$, $\Pi = 565.28$
- $M = 130$: $\vec{x} = (50, 39, 31, 10)$, $\vec{z} = (52, 39, 26, 13)$, $\Pi = 438.74$

Даље добијамо $y = 860.72 - 3.29 \cdot x$ за регресију. ❑

Понављајући ове програме много пута приметићемо оно што се види и из само ова два примера, да коефицијент нагиба регресионе линије, у првом примеру позитиван број ($b = 7$) каже нам да се површине повећавају, а у другом негативан ($b = -3.29$) каже нам да се оне смањују, упозорава на можда дозвољена статистичка одступања око нулте вредности. У том случају коефицијент пресека иначе удаљене ординате ($\vec{a}$, био је -505 и 861) очекивано је велик број.

Пример 3. Погледајмо случајне исходе сличне претходним, али са већим бројем опита.

X	***Y***	***X²***	***XY***
1000	12 590	1000 000	12 590 000
1010	12 288	1 020 100	12 410 880
1020	11 274	1 040 400	11 499 480
1030	14 421	1 060 900	14 853 630
Σx = 4060	Σy = 50 573	Σx² = 4 121 400	Σxy = 51 353 990

$$y = -32\,819 + 45 \cdot x$$

X:	1000	1010	1020	1030
Y:	12 181	12 631	13 081	13 531

Истим кораком $\Delta M = 10$, али са приближно десет пута већим бројевима опита, површине расту брже од шест пута. У доњем реду (*Y*:) су процењене вредности регресијом. ❑

Ова „шест пута“ већа брзина пораста површине могла би бити успоравање приближавања $\vec{x} \to \vec{z}$ (са порастом броја понављања опита M), али и не мора, јер све ове промене коефицијента нагиба регресије (b) су у границама статистичке грешке. Тако мала промена нагиба праве која би требала бити хоризонтална (паралелна са апсцисом), негде на далеким местима може пресецати ординату на значајно већој висини или дубини ($\vec{a}$). То потврђују понављања овог примера, од којих наводим само један случај.

Пример 4. Са истим почетним условима претходног примера, генерисани су нови бројеви:

X	*Y*	*X*²	*XY*
1000	16 902.96	1000 000	16 902 960.0
1010	19 460.27	1 020 100	19 654 872.7
1020	12 782.19	1 040 400	13 037 833.8
1030	4 431.39	1 060 900	4 564 331.7
Σx = 4060	Σy = 53 576.81	Σx² = 4 121 400	Σxy = 54 159 998.2

$$y = 460\ 935.76 - 440.93 \cdot x$$

X:	1000	1010	1020	1030
Y:	19 936	15 526	11 116	6 706

Површине сада опадају, наизглед прилично брзо ($\vec{b} = -441$), али то је у границама статистичке толеранције за релативно велике бројеве понављања опита (прва колона). Отуда и тако висока тачка пресека ($\vec{a} = 460\ 936$) праве линије регресије са веома удаљеном ординатом. ❑

У аналогији са физичком силом, нетолерисано успоравање приближавања двају вектора, $\vec{x}$ и $\vec{z}$, представљало би не само уобичајено опадање дејства на већој удаљености (Confidential), него и „нестабилност" самих носиоца поља, рецимо њиховим краћим веком живота – који фотони или гравитони немају. Опет напомињем, ако би одступања ових примера била статистички значајна.

Подсећам, за дати вектор случајних бројева $\vec{x}$, односно дати број корака M, већа површина Π је показатељ већег угла између тог и вектора $\vec{z}$ који има правац вероватноће $\vec{p}$. Повећање размере дужина вектора наспрам угла између њих указује на слабију „адаптацију" првог вектора другом. Сам коефицијент b, је однос пораста површине и прираста повољних исхода ($b = \Delta y : \Delta x$), нагиб линије регресије утолико већи што релативне вероватноће спорије конвергирају стварним.

Пример 5. Расподела коју смо задали у претходним примерима (0.4, 0.3, 0.2, 0.1) нема претежну, неку вероватноћу која је већа од збира осталих (Channel, Теорема), овде [9.]. То није тако рецимо са расподелом $\vec{p} = (0.6, 0.2, 0.1, 0.1)$ за коју генеришемо следеће исходе и регресију:

X:	*M* = 100	*M* = 110	*M* = 120	*M* = 130
Y:	*Π* = 489	*Π* = 171	*Π* = 405	*Π* = 511

$$y = 399 - 0.04x$$

Регресија је скоро нулта, површине су приближно константне. ❑

Доминантна једна од вероватноћа расподеле могла би потећи од осталих исхода толико ређих да су вектори $\vec{x}$ и $\vec{z}$ чешће усклађени, да већи коефицијент првог одговара већем дугог и обрнуто, те је скаларни производ $\vec{x} \cdot \vec{z}$ већ у старту максималан, а угао $\varphi = \sphericalangle(\vec{x}, \vec{z})$ минималан, као и синус тог угла, па је онда и тражена површина мања.

Понављањем оваквих „извеснијих" расподела може се чинити да је учесталост тих „скоро нултих" коефицијената нагиба ($\vec{b}$) већа од претходних примера, али процене значајности нису ме убедиле. За сада немам јасне одлуке и оставићу то питање отворено, за друге прилике. Ништа се значајно не мења нити са повећањем броја опита, што демонстрирам следећим примером.

Пример 6. Настављамо са вероватноћама расподеле $\vec{p} = (0.6, 0.2, 0.1, 0.1)$, са већим бројевима поновљених опита $M \geq 1000$, такође константног корака $\Delta M = 10$.

X:	*M* = 1000	*M* = 1010	*M* = 1020	*M* = 1030
Y:	*П* = 14 340	*П* = 19 319	*П* = 21 226	*П* = 14 252

$$y = 17\ 287 - 0.003x$$

Регресија је још стабилније паралелна линија. ❑

Код боље изјашњених расподела вероватноћа, као у примерима 5 и 6, нагомилавање позитивних исхода као да показује (незнатно) већу аналогију са Кеплеровим другим законом од дифузнијих, какви су у примерима 1 и 2. Међутим, могло би то бити недовољно статистички значајно.

Пример 7. Вероватноћа је веома дифузна, уједначена $\vec{p} = (0.27, 0.26, 0.24, 0.23)$, са резултатима:

X:	*M* = 100	*M* = 110	*M* = 120	*M* = 130
Y:	*П* = 535	*П* = 299	*П* = 618	*П* = 436
y:	469	471	473	476

$$y = 446.70 + 0.22x$$

У доњем реду, сенченом, наведене су процене регресијом. ❑

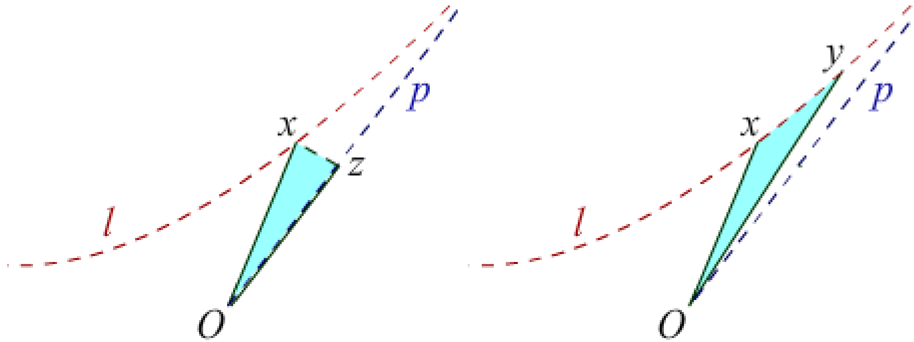
Дакле, у екстремним случајевима, како извесности тако и неизвесности исхода, дате расподеле вероватноћа, мог рачунара, показују веома приближно константне површине, $y = П(x)$, између вектора $\vec{x}$ (датим вероватноћама генерисаних бројева повољних исхода) и очекиваних $\vec{z}$ бројева исхода. Другим речима, у њиховом случају такође постоји корелација аналогије Кеплеровог другог закона са оваквим нагомилавањима повољних исхода.

Приметимо да „статистичка корелација“ није исто што „узрочно-последичне везе“, али да то сада није толико битно као у „свакодневном животу“, јер овде она открива аналогију. Аналогија о којој причамо је „једнакост“ попут исте форме сабирања рецимо јабука и метара, који нису физикално исте мере, нити су чак ни апстрактно исти појмови. Међутим, очекујем даље да постоје још неке за сада тајне везе између физичких сила и „сила“ којима природа спонтано, кажем радије, реализује вероватније исходе, а тиме и мање информативна стања.

Оно што је ипак различито у случајевима статистичке и прецизне науке је нагомилавање грешака. Када радимо са великом толеранцијом „тачности“, уз статистичку оцену ваљаности, појавиће се и увећавање грешака нетипично за прецизно рачунање. Свако ко је на компјутеру додавао сабирке рецимо развоја синусне функције у Тејлоров ред зна о чему говорим. Ти бројеви су заокружени на 5, 10, 15, или колико год хоћете децимала, али садрже локално занемарљиве нетачности које због нагомилавања збирно постају веће од Тејлорове процене за остатак реда.

24. Примери регресије II

У претходним „примерима регресије“ радили смо по шеми следеће слике лево, а у следећим радићемо по десној. Дата је расподела вероватноћа $\vec{p} = (p_1, p_2, \dots, p_n)$, у неком $n = 1, 2, 3, \dots$ димензионалном простору случајних догађаја $\omega_1, \omega_2, \dots, \omega_n \in \Omega$. Понављамо опит $M = 1, 2, 3, \dots$ пута и бележимо ξ_j број успешних реализација j-тог догађаја, иначе вероватноће p_j, где индекси имају вредности $j = 1, 2, \dots, n$. Тако се понављањем формира вектор $\vec{x} = (\xi_1, \xi_2, \dots, \xi_n)$ све веће дужине (интензитета) $\|\vec{x}\| = \sqrt{\xi_1^2 + \xi_2^2 + \dots + \xi_n^2}$. У свим примерима $n = 4$.



Очекиване вредности координата вектора $\vec{x}$ након M опита су $\zeta_j = Mp_j$. Са њима дефинишемо вектор $\vec{z} = (\zeta_1, \zeta_2, \dots, \zeta_n)$ интензитета $\|\vec{z}\| = \sqrt{\zeta_1^2 + \zeta_2^2 + \dots + \zeta_n^2} = M\sqrt{p_1^2 + p_2^2 + \dots + p_n^2} = M\vec{p}$, који је увек истог правца p, али растућег интензитета (са бројем M понављања опита), приказан сликом лево. Иако су координате вектора $\vec{x}$ случајне варијабле, лако је разумети да угао између два вектора, $\varphi = \sphericalangle(\vec{x}, \vec{z})$, опада, смањује се синус тог угла, а „неочекиван“ резултат примера је откриће, или потврда да је површина, $\Pi = \|\vec{x}\|\|\vec{z}\| \sin\varphi$, паралелограма којег вектори разапињу константна у статистичком смислу. Шрафирани троугао је половине површине паралелограма.

У наставку, уместо вектором $\vec{z}$, константног правца и интензитета растућег упоредо са вектором $\vec{x}$, бавимо се вектором $\vec{y} = (\eta_1, \eta_2, \dots, \eta_n)$ интензитета $\|\vec{y}\| = \sqrt{\eta_1^2 + \eta_2^2 + \dots + \eta_n^2}$ за корак ΔM већег од првог, $\|\vec{x}\| < \|\vec{y}\|$. Ова два вектора разапињу паралелограм неке површине $\Pi = \|\vec{x}\|\|\vec{y}\| \sin\varphi$ такође, а промењљивост те површине је предмет наредних проверавања.

Пример 1. Задата вероватноћа је $\vec{p} = (0.4, 0.3, 0.2, 0.1)$, почетни број понављања $M = 100$, са кораком $\Delta M = 10$ и генерисана су четири пара ($K = 4$) упоредних низова исхода $\vec{x}$ и $\vec{y}$:

- $M_x = 100$, $\vec{x} = (36, 24, 29, 11)$, $M_y = 110$, $\vec{y} = (42, 39, 16, 13)$, $\Pi = 1\,086.79$
- $M_x = 110$, $\vec{x} = (40, 35, 20, 15)$, $M_y = 120$, $\vec{y} = (46, 31, 27, 16)$, $\Pi = 535.61$
- $M_x = 120$, $\vec{x} = (47, 36, 26, 11)$, $M_y = 130$, $\vec{y} = (46, 46, 27, 11)$, $\Pi = 569.11$
- $M_x = 130$, $\vec{x} = (59, 38, 26, 7)$, $M_y = 130$, $\vec{y} = (70, 28, 29, 13)$, $\Pi = 1\,162.57$

На основу два генерисана низа случајних бројева рачун даје даље, рецимо за четврти корак:

$$\vec{x} \cdot \vec{y} = 6\,039,\ \|\vec{x}\| = 75.17,\ \|\vec{y}\| = 81.82,\ \cos\varphi = 0.98917,\ \sin\varphi = 0.18904.$$

Након сва четири корака налазимо параметре линеарне регресије:

$$\Sigma x = 460.00,\ \Sigma y = 3\,354.08,\ \Sigma x^2 = 53\,400.00,\ \Sigma xy = 387\,023.42$$

и коефицијенте линеарне регресије:

$$a = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{K(\Sigma x^2) - (\Sigma x)^2} = 538.54300,\ b = \frac{K(\Sigma xy) - (\Sigma x)(\Sigma xy)}{K(\Sigma x^2) - (\Sigma x)^2} = 2.60850,$$

$$y = a + bx.$$

Регресиона права линија скоро је хоризонтална ($b = 2.6$), али је релативно високо (просечне висине, површине $y =$ 838.52), па на већој висини ($a = 538.5$) пресеца ординату. ❑

Паралелност линије регресије са апсцисом (x-осом) говори о константности ординате (вредности y-а) која је овде мера површине. Статистички је скоро немогућ случај тачне паралелности ($b = 0$), јер су ово значајно случајни процеси, па „паралелност“ укључује толеранцију статистичке грешке, која је иначе променљива величина. Зато бележим још један претходни поступак, од многих.

Пример 2. Исти су услови претходног примера, али не њима и генерисани бројеви:

- $M_x = 100$, $\vec{x} =$ (47, 31, 13, 9), $M_y =$ 110, $\vec{y} =$ (40, 37, 23, 10), $\Pi =$ 797.56
- $M_x = 110$, $\vec{x} =$ (35, 30, 29, 16), $M_y =$ 120, $\vec{y} =$ (44, 46, 18, 12), $\Pi =$ 1165.11
- $M_x = 120$, $\vec{x} =$ (48, 29, 31, 12), $M_y =$ 130, $\vec{y} =$ (56, 40, 23, 11), $\Pi =$ 930.73
- $M_x = 130$, $\vec{x} =$ (49, 47, 21, 13), $M_y =$ 140, $\vec{y} =$ (50, 47, 30, 13), $\Pi =$ 609,42

Коефицијенти, $a =$ 1 794.33 и $b =$ -7.99, регресије $y = a + bx$, сада процењују да се ради о (невеликом) опадању површине. ❑

Са приближно $M = 100$ поновљених опита статистика толерише $b = 0 \pm 40$ као „хоризонталан“ нагиб линије регресије, па ипак, или да кажемо баш због тога, са повећањем броја опита на око $M = 1000$ толеранција се подиже на $b = 0 \pm 400$. При томе треба знати да је та „толеранција“ и релативан и променљив појам.

Пример 3. Исти су услови претходног примера, али не и њима генерисани бројеви:

- $M_x = 1000$, $\vec{x} =$ (419, 282, 205, 94), $M_y =$ 1010, $\vec{y} =$ (408, 314, 184, 104), $\Pi =$ 22 687.50
- $M_x = 1010$, $\vec{x} =$ (422, 308, 187, 93), $M_y =$ 1020, $\vec{y} =$ (406, 301, 213, 100), $\Pi =$ 17 735.19
- $M_x = 1020$, $\vec{x} =$ (398, 301, 209, 112), $M_y =$ 1030, $\vec{y} =$ (419, 298, 213, 100), $\Pi =$ 11 744.34
- $M_x = 1030$, $\vec{x} =$ (415, 314, 181, 120), $M_y =$ 1040, $\vec{y} =$ (407, 319, 201, 113), $\Pi =$ 13 038.92

Коефицијенти, $a =$ 370 907.77 и $b =$ -349.37, регресије $y = a + bx$, сада осцилују углавном унутар опсега $b = 0 \pm 400$. ❑

Очекивање $\mu = E(\vec{x})$ и варијанса $\sigma^2 = \mathrm{Var}(\vec{x})$ су две важне специфичности низа вероватноћа. Као што знамо, очекивање је просек, а варијанса је очекивање квадрата одступања од просека:

$$E(\vec{x}) = \sum_{j=1}^{n} \xi_j p_j,\quad \mathrm{Var}(\vec{x}) = E[(\vec{x} - \mu)^2] = \sum_{j=1}^{n} \left(\xi_j - \mu\right)^2 p_j$$

За унапред дате бројеве понављања опита, M, дефинисали смо и $E\left(\xi_j\right) = Mp_j = \zeta_j$, па онда и $\vec{z} = Mp_j$. Те величине су низови очекивања, за разлику од збирног очекивања, броја μ. Такође, постоје различите варијабилности, поред просте σ^2, такође важне за разумевање површина о којима говоримо.

Пример 4. Задата је нова вероватноћа $\vec{p} = (0.6, 0.2, 0.1, 0.1)$, стари број понављања $M = 100$ и истим кораком $\Delta M = 10$ генерисана су четири пара ($K = 4$) упоредних низова исхода $\vec{x}$ и $\vec{y}$:

- $M_x = 100$, $\vec{x} =$ (57, 25, 8, 10), $M_y = 110$, $\vec{y} =$ (76, 16, 10, 8), Π =1 042.97
- $M_x = 110$, $\vec{x} =$ (64, 25, 10, 11), $M_y = 120$, $\vec{y} =$ (69, 26, 15, 10), Π = 330.95
- $M_x = 120$, $\vec{x} =$ (62, 25, 12, 21), $M_y = 130$, $\vec{y} =$ (85, 21, 8, 16), Π =1259.25
- $M_x = 130$, $\vec{x} =$ (81, 26, 15, 8), $M_y = 130$, $\vec{y} =$ (84, 25, 12, 19), Π = 993.00

Отуда за регресију, $y = a + bx$, израчунавамо коефицијенте $a = 11.42$ и $b = 7.78$. ❑

Наводим још случајних бројева које ми сада рутински избацује компјутер, а на које ћу се касније моћи једноставније позивати.

Пример 5. Исти су услови као у претходном примеру, а генерисана су нова четири пара низова:

- $M_x = 100$, $\vec{x} =$ (73, 12, 6, 9), $M_y = 110$, $\vec{y} =$ (64, 25, 11, 10), Π =1 152.99
- $M_x = 110$, $\vec{x} =$ (58, 22, 15, 15), $M_y = 120$, $\vec{y} =$ (79, 18, 11, 12), Π = 1 010.45
- $M_x = 120$, $\vec{x} =$ (61, 26, 16, 17), $M_y = 130$, $\vec{y} =$ (71, 30, 15, 14), Π =451.76
- $M_x = 130$, $\vec{x} =$ (73, 26, 20, 11), $M_y = 130$, $\vec{y} =$ (76, 38, 11, 15), Π = 1 214.69

За регресију, $y = a + bx$, добијамо нове коефицијенте $a = 1\,387.12$ и $b = -3.74$. ❑

Одговарајући пример дифузној вероватноћи претходног наслова [23. Пример 6], следећи је.

Пример 6. Задате су вероватноће, $\vec{p} = (0.27, 0.26, 0.24, 0.23)$ уз остале исте услове из претходног примера, са новим генерисаним паровима случајних низова:

- $M_x = 100$, $\vec{x} =$ (32, 24, 20, 24), $M_y = 110$, $\vec{y} =$ (33, 26, 31, 20), Π = 569.36
- $M_x = 110$, $\vec{x} =$ (37, 23, 19, 31), $M_y = 120$, $\vec{y} =$ (30, 34, 29, 27), Π = 957.77
- $M_x = 120$, $\vec{x} =$ (30, 27, 32, 31), $M_y = 130$, $\vec{y} =$ (37, 36, 29, 28), Π = 628.04
- $M_x = 130$, $\vec{x} =$ (42, 38, 29, 21), $M_y = 130$, $\vec{y} =$ (39, 36, 32, 33), Π = 852.39

За регресију, $y = a + bx$, налазимо коефицијенте $a = 106.04$ и $b = 5.73$. ❑

Посебно ћемо се бавити дисперзијама ових површина зависно од броја опита M и од расутости самих задатих вероватноћа $\vec{p}$, јер их је тешко разликовати на први поглед.

25. Средње вредности

Из алгебре познајемо различите средње вредности низа $n = 1, 2, 3, \dots$ бројева $\vec{x} = (x_1, x_2, \dots, x_n)$, као што су хармонијска, геометријска и аритметичка средина:

$$H(\vec{x}) = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}},\ G(\vec{x}) = \sqrt[n]{x_1 x_2 \dots x_n},\ A(\vec{x}) = \frac{x_1+x_2+\cdots+x_n}{n}.$$

За њих за дате бројеве увек важе неједнакости

$$\min_{1\le j\le n}\{x_j\} \le H \le G \le A \le \max_{1\le j\le n}\{x_j\}$$

где једнакост стоји ако и само ако су сви ови бројеви (x_j) међусобно једнаки.

У статистици су средња вредност (mean), медијана (median) и мод (mode) три главна начина за означавање просечних вредности листе бројева. Први је поменута аритметичка средина, која се често подразумева под просеком, а налази као горе, тако што се бројеви саберу и збир подели бројем сабирака. Медијан је средњи број на листи пореданој од најмањег до највећег. Најчешће појављивана вредност на листи је мод.

Математичко очекивање везујемо за вероватноће расподеле, дакле за бојеве $p_j \in (0,1)$ такве да је $p_1 + p_2 + \cdots + p_n = 1$, појединих случајних варијабли $X = \{x_1,\ x_2, \dots, x_n\}$. Елементи овог скупа обично су реални бројеви ($x_j \in \mathbb{R}$), често и комплексни ($\mathbb{C}$), али смо за сада разматрали углавном природне ($\mathbb{N}$). Дискретно математичко очекивање таквих, дефинише

$$E(X) = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n$$

при чему је дозвољено да $n \to \infty$, ако овај збир конвергира. За непрекидну случајну промењљиву X, са густинама вероватноће $\rho(x)$ из домена $x \in (a, b)$, где $\int_a^b \rho(x)\, dx = 1$, ово очекивање је

$$E(X) = \int_a^b x\rho(x)\ dx$$

када несвојствени интеграл $\int_a^b |x|\rho(x)\, dx$ конвергира. У оба случаја, дискретном и непрекидном, математичко очекивање означавамо и грчким словом μ, пазећи да избегнемо забуне.

Пример 1. У претходним примерима регресије математичка очекивања координата вектора чине вектор, $E(\vec{x}) = E(\xi_1, \xi_2, \dots, \xi_n) = (Mp_1, Mp_2, \dots, Mp_n) = M(p_1, p_2, \dots, p_n) = M\vec{p}$, где је M био број опита. Очекивање одговарајућег скупа био би број $E(X) = (\xi_1 + \xi_2 + \cdots + \xi_n)/n$. ❑

Пример 2. Расподела дискретних вероватноћа $\vec{p}$ дефинише низ информација $h_j = -\log p_j$ са математичким очекивањем $E(\vec{h}) = \sum_{j=1}^{n} h_j p_j$. То је Шенонова информација. ❑

Пример 3. Случајне варијабле скупа X све једнаке неком броју, константи c, имаће том броју једнако (математичко) очекивање, $E(c) = c$. Све случајне варијабле скупа X множене бројем c, даће очекивање множено истим бројем, $E(cX) = cE(X)$. Очекивање збира варијабли једнако је збиру очекивања, $E(X + Y) = E(X) + E(Y)$. Очекивање производа независних промењљивих је производ очекивања, $E(XY) = E(X)E(Y)$. То су добро познате особине очекивања. ❑

Варијанса је (математичко) очекивање квадратних одступања од $\mu = E(X)$:

$$\mathrm{Var}(X) = \frac{1}{n}\sum_{j=1}^{n}(x_j - \mu)^2 = E[(X-\mu)^2] = E(X^2) - \mu^2$$

где X^2 означава скуп квадрираних елемената скупа X. Варијансу означавамо и са σ^2 и она је једна од најчешћих дисперзија, расипања случајне промењљиве око средње вредности, па се често тако и назива – дисперзијом. Особине варијансе су такође добро познате.

Пример 4. $\sigma^2(X) = 0$ ако и само ако скоро све[5] варијабле имају исту вредност. $\sigma^2(X + c) = \sigma^2(X)$. $\sigma^2(cX) = c^2\sigma^2(X)$. За независне важи $\sigma^2(X + Y) = \sigma^2(X) + \sigma^2(Y)$. Функција $E(X - \lambda)^2$ по λ има минимум за $\lambda = E(X)$, тј. $\sigma^2(X) \le E(X - \lambda)^2$ за сваки број λ, са једнакошћу само за минимум. ❑

Коваријанса је очекивање производа одступања од очекивања двеју случајних варијабли:

$$\mathrm{Cov}[(X - \mu_x))(Y - \mu_y)] = E(XY) - \mu_x\mu_y$$

где стоји $\mu_x = E(X)$ и $\mu_y = E(Y)$. Када веће вредности прве варијабле нагињу ка већим друге, тада је $(X - \mu_x))(Y - \mu_y)$ позитиван број и кажемо да су те две варијабле позитивне корелације, иначе је негативан и корелација варијабли је негативна.

Особине коваријансе су, такође, добро познате:

1. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$;
2. $\mathrm{Cov}(X, Y) = 0$, за независне варијабле;
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$;
4. $\mathrm{Cov}(\lambda X, Y) = \lambda\mathrm{Cov}(X, Y)$;
5. $\mathrm{Cov}(X + c, Y) = \mathrm{Cov}(X, Y)$;
6. $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(Y, Z)$.

Коефицијент корелације је норма коваријансе

$$\rho_{xy} = \mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

са, добро познатим, особинама:

1. $-1 \le \mathrm{Corr}(X, Y) \le 1$;
2. Ако је $\mathrm{Corr}(X, Y) = 1$, тада је $Y = aX + b$, где је $a > 0$;
3. Ако је $\mathrm{Corr}(X, Y) = -1$, тада је $Y = aX + b$, где је $a < 0$;
4. $\mathrm{Corr}(aX + b, cY + d) = \mathrm{Corr}(X, Y)$ за $a, b > 0$.

[5] скоро све – све сем оних нулте верованоће

26. Дисперзије

Средње вредности, тачније математичка очекивања су и неке од дисперзија, просечна одступања од неких важних вредности. У претходном наслову бавили смо се само дефиницијама и њиховим шире познатим особинама, али не и неким важнијим последицама. Али постоје и дисперзије које се не заснивају на средњим вредностима тамо поменутим, а поједине једва да литература уопште и помиње.

Теорема 1. (Чебишљева неједнакост) Ако постоји $E(X^2)$, очекивање квадрата случајне варијабле X, тада за свако $\varepsilon > 0$ важи неједнакост

$$P\{|X| \geq \varepsilon\} \leq \frac{E(X^2)}{\varepsilon^2}$$

која каже да је вероватноћа да је варијабла већа од епсилон мања од очекивања квадрата варијабле подељеним са квадратом епсилона.

Доказ: Укратко, $E(X^2) = \int_{-\infty}^{+\infty} x^2 \rho(x)\, dx = \int_{|x|\geq\varepsilon} x^2 \rho(x)\, dx + \int_{|x|<\varepsilon} x^2 \rho(x)\, dx \geq$

$$\int_{|x|\geq\varepsilon} x^2 \rho(x)\, dx \geq \int_{|x|\geq\varepsilon} \varepsilon^2 \rho(x)\, dx = \varepsilon^2 \int_{|x|\geq\varepsilon} x^2 \rho(x)\, dx \geq \varepsilon^2 P\{|X| \geq \varepsilon\}. \blacksquare$$

Када уместо X ставимо $X - \mu$ Чебишљева неједнакост постаје $P\{|X - \mu| \geq \varepsilon\} \leq \sigma^2(X)/\varepsilon^2$, где је μ математичко очекивање. Ево како она може послужити за доказе једне од врста закона великих бројева.

Посматрајмо неки случајни догађај из $\omega \in \Omega$. Нека је ω_k догађај да се у k-том понављању опита реализује дати ω, а $\overline{\omega}_k$ да се не реализује. Вероватноћа реализације је $P(\omega_k) = p$, док је супротна вероватноћа $P(\overline{\omega}_k) = 1 - p = q$. Индикатор дешавања $I_k = 1$ ако се догађај десио и $I_k = 0$ ако се догађај није десио. Отуда је број реализација датог догађаја $\xi = I_1 + I_2 + \cdots + I_M$ у M понављања. Због независности описаних понављања, независне варијабле су и ови индикатори, па је

$$\sigma^2(\xi) = \sigma^2(I_1) + \sigma^2(I_2) + \cdots + \sigma^2(I_M)$$

Очекивање индикатора је $E(I_k) = 1 \cdot P(\omega_k) + 0 \cdot P(\overline{\omega}_k) = p$. Због $I_k^2 = I_k$, варијанса индикатора је $\sigma^2(I_k) = E(I_k^2) - [E(I_k)]^2 = E(I_k) - p^2 = p - p^2 = pq$. Како је $P(\omega_k) = p$, за свако $k = 1, 2, \ldots, M$, јер ω има исту вероватноћу у сваком поновљеном опиту, то је $\sigma^2(I_k) = P(\omega_k) - [P(\omega_k)]^2 = pq$ и $\sigma^2(\xi) = Mpq$. Ово су кључни резултати теорије вероватноће, а такође потребан додатак прошлим примерима регресије и зато толико детаљишем.

Пример 1. Уочимо аритметичку средину индикатора $A_M(I) = \frac{1}{M}(I_1 + I_2 + \cdots + I_M)$. Нека је $E(A_M) = a$, па је $\sigma^2(A_k) = \frac{1}{M^2}\sum_{k=1}^{M} \sigma^2(I_k) = \frac{\sigma^2}{M}$, а из Чебишљеве неједнакости

$$P\{|A_M - a| \geq \varepsilon\} \leq \frac{\sigma^2(A_M)}{\sigma^2} = \frac{\sigma^2}{M\varepsilon^2}$$

Овај број тежи нули када $n \to \infty$, а то је Бернулијев закон великих бројева. ❑

Сличан, мало општији резултат је следећи, када не посматрамо само индикаторе.

Пример 2. Нека су $X_1, X_2, X_3, \ldots$ независне случајне промењливе са истом расподелом, сталним очекивањем $E(X_k) = \mu$ и константном коначном варијансом $\sigma^2(X_k) = \sigma^2$, за свако k. Тада њихова аритметичка средина $\bar{X}_n(X) = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$ тежи ка μ, када $n \to \infty$. На исти претходни начин, сада из $E(\bar{X}_n) = \mu$, $\sigma^2(\bar{X}_n) = \sigma^2/n$ и опет Чебишљеве неједнакости добијамо

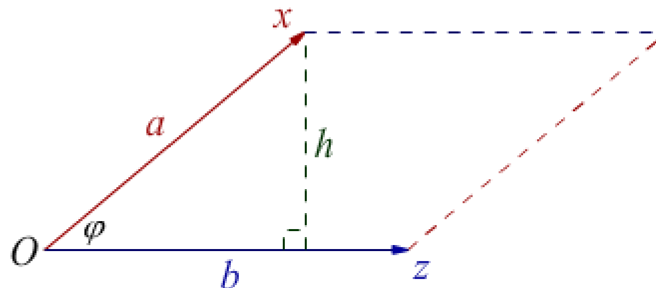
$$P\{|\bar{X}_n - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{n\varepsilon^2}$$

када $n \to \infty$. То је поопштење претходног Бернулијевог закона великих бројева. ❑

Ни овај закон великих бројева, наравно, није последња реч теорије вероватноће. Сличне овоме разумемо да аритметичка средина случајних промењљивих истог очекивања μ губи случајност у смислу да је A_M све ближе тој очекиваној вредности када је број понављања M већи. Вероватноћа сваке мале разлике ове аритметичке средине и константног очекивања тежи нули, $|A_M - \mu| \to 0$, са неограниченим понављањем опита (када $M \to \infty$).

Закони великих бројева уопште откривају нам конвергенције вероватноћа, њихове услове и смер, али неће се бавити детаљима који нас при томе могу заинтересовати. Дисперзјама се зато бавимо и на друге начине, а један од њих који користим у текстовима скрипте је визуелизација ових иначе веома апстрактних проблема.

Основни су скаларни и псеудо скаларни производ вектора, типични за информацију перцепције, а око случајних варијабли они су скоро свеприсутне величине. Математичка статистика као да их за сада потцењује, па ћемо их још мало експлоатисати. Први говоре о производу дужина вектора и косинуса угла између њих, а други о производу тих дужина и синуса угла, а дисперзије укључују и налажење коефицијената који ће ове производе стабилизовати, нормирати.



Површина паралелограма којег разапињу вектори $\vec{x}$ и $\vec{z}$ дужина $a = \|\vec{x}\|$ и $b = \|\vec{z}\|$ је $\Pi = ab\sin\varphi$, где је $\varphi = \sphericalangle(\vec{z}, \vec{x})$ угао између вектора. Наиме, површина паралелограма је производ основице и висине на њу $\Pi = bh$, а из правоуглог троугла коју гради та висина као катета и страница лево као хипотенуза, $h = a\sin\varphi$, сменом добијамо поменути резултат. Добијамо „синусну дисперзију“.

Остаје нам само да бирамо дефиницију дужине вектора, не мењајући му правац и смер, тако да за што је могуће више расподела вероватноћа увек добијамо неку површину која ће остајати иста без обзира на број понављања опита. Један предлог изнећу у следећем наслову.

27. Сила неизвесности

Потребу за корекцијом „површине“ паралелограма којег разапињу вектори $\vec{x}$ и $\vec{z}$ намеће слобода дефиниције „дужине“ тих вектора [23.]. Ако у примерима регресије, дужину вектора не сматрамо простим збиром реализованих исхода, него рецимо њиховим кореном, регресија постаје мирнија и њена хоризонтална права стабилнија. У том случају [23. Пример 1] добија следећи исход.

Пример 1. За вероватноће поставимо $\vec{p} = (0.4, 0.3, 0.2, 0.1)$ и генеришимо случајне бројеве почев од $M = 100$, кораком $\Delta M = 10$. Уз свако M, добијамо збирове исхода $\vec{x}$ и очекивања $\vec{z}$, али за дужине тих вектора узмимо корене тих збирова. Тада израчунавамо површину коју вектори разапињу:

- $M = 100$: $\vec{x} = (42, 31, 19, 8)$, $\vec{z} = (40, 30, 20, 10)$, $\Pi = 2.86$
- $M = 110$: $\vec{x} = (34, 36, 28, 12)$, $\vec{z} = (44, 33, 22, 11)$, $\Pi = 11.84$
- $M = 120$: $\vec{x} = (42, 39, 27, 12)$, $\vec{z} = (48, 36, 24, 12)$, $\Pi = 7.20$
- $M = 130$: $\vec{x} = (58, 38, 26, 8)$, $\vec{z} = (52, 39, 26, 13)$, $\Pi = 7.15$

Линија регресије $y = a + bx$, где апсциса M дефинише ординату Π, има коефицијенте $a = -2.24$ и $b = 0.08$, дакле много је паралелнија права апсциси (x-оси) од претходне. ❑

Да се заиста добијају значајно стабилније површине, који разапињу вектори $\vec{x}$ и $\vec{z}$ истих праваца а краћих интензитета (корена броја реализација у M покушаја), показује вишеструко генерисање тих података, као у следећем примеру.

Пример 2. Расподела вероватноћа је и даље $\vec{p} = (0.4, 0.3, 0.2, 0.1)$, почетно $M = 100$ једнако као у претходном примеру, али је корак десет пута већи $\Delta M = 100$.

- $M = 100$: $\vec{x} = (46, 24, 16, 14)$, $\vec{z} = (40, 30, 20, 10)$, $\Pi = 54.77$
- $M = 200$: $\vec{x} = (84, 58, 43, 15)$, $\vec{z} = (80, 60, 40, 20)$, $\Pi = 7.00$
- $M = 300$: $\vec{x} = (127, 79, 59, 35)$, $\vec{z} = (120, 90, 60, 30)$, $\Pi = 164.32$
- $M = 400$: $\vec{x} = (161, 116, 91, 32)$, $\vec{z} = (160, 120, 80, 40)$, $\Pi = 14.12$

Линија регресије ($y = a + bx$) има коефицијенте $a = 6.51$ и $b = 0.02$, дакле веома хоризонтална је. ❑

Осцилације површине до 20 пута говоре о великој дисперзији (случајности) ових варијабли, али и помало нестварна стабилност регресије ($b = 0.02$), односно непромењљивост средње вредности површине, о неочекивано снажној аналогији апстрактних „сила неизвесности“ са силама небеске механике, односно Кеплеровим другим законом.

Пример 3. Расподела вероватноћа је нова $\vec{p} = (0.6, 0.2, 0.1, 0.1)$, почетно $M = 1000$ увећано је десет пута, а корак остаје једнак претходном $\Delta M = 100$.

- $M = 1000$: $\vec{x} = (581, 218, 95, 106)$, $\vec{z} = (600, 200, 100, 100)$, $\Pi = 24.07$
- $M = 1100$: $\vec{x} = (655, 225, 109, 111)$, $\vec{z} = (660, 220, 110, 110)$, $\Pi = 6.53$
- $M = 1200$: $\vec{x} = (708, 276, 121, 95)$, $\vec{z} = (720, 240, 120, 120)$, $\Pi = 45.37$
- $M = 1300$: $\vec{x} = (803, 265, 113, 119)$, $\vec{z} = (780, 260, 130, 130)$, $\Pi = 24.65$

Линија регресије ($y = a + bx$) има коефицијенте $a = -18.76$ и $b = 0.04$, хоризонтална је веома и даље. ❑

Ово су биле „ангажоване“ вероватноће расподеле, са једном већом од збира осталих, али нема значајних разлика нити са „разблаженим“, које су све четири око четвртине. То имамо у следећем примеру.

Пример 4. Расподела вероватноћа је нова $\vec{p} = (0.27, 0.26, 0.24, 0.23)$, почетно $M = 10\,000$ увећано је опет десет пута, колико и корак $\Delta M = 1000$.

- $M = 10\,000$: $\vec{x} = (2662, 2604, 2431, 2303)$, $\vec{z} = (2700, 2600, 2400, 2300)$, $\Pi = 49.26$
- $M = 11\,000$: $\vec{x} = (2962, 2858, 2726, 2454)$, $\vec{z} = (2970, 2860, 2640, 2530)$, $\Pi = 115.04$
- $M = 12\,000$: $\vec{x} = (3113, 3182, 2833, 2872)$, $\vec{z} = (3240, 3120, 2880, 2760)$, $\Pi = 186.27$
- $M = 13\,000$: $\vec{x} = (3455, 3377, 3145, 3023)$, $\vec{z} = (3510, 3380, 3120, 3990)$, $\Pi = 68.80$

Линија регресије ($y = a + bx$) има коефицијенте $a = -44.52$ и $b = 0.01$ и такође је хоризонтална попут оних у претходна три примера. ❑

Синусна дисперзија се показује знатно прецизнијом када уместо простог збира реализација мере интензитета два вектора у производу површине употребимо корене бројева тих исхода. Ако број приближно $\Pi = (Mp)^2 \sin\varphi$, заменимо бројем $\Pi = (Mp)\sin\varphi$, услов Π = const. постаје тачнији. То ће се поновити и у примерима два вектора истих корака ΔM разлике у бројевима опита M, а да је то тако видимо прво помоћу закона великих бројева [26. Пример 1 и 2], а затим и из наредних примера.

Пример 5. Расподела је из претходног примера $\vec{p} = (0.27, 0.26, 0.24, 0.23)$, почетно $M = 10\,000$, са корацима $\Delta M = 1000$, али вектори су $\vec{x}$ и $\vec{y}$ из примера регресије [24.], када је други за корак испред првог.

M	$\vec{x}$	$\vec{y}$	Π
10 000	(2664, 2635, 2385, 2316)	(2965, 2854, 2675, 2506)	82.90
11 000	(2975, 2815, 2663, 2547)	(3190, 3089, 2858, 2863)	108.07
12 000	(3263, 3079, 2859, 2799)	(3511, 3368, 3188, 2933)	134.86
13 000	(3526, 3373, 3136, 2965)	(3755, 3687, 3326, 3232)	90.89

Регресија ($y = a + bx$) има коефицијенте a = 45.79062 и b = 0.00508. ❑

Површине су приближно једнаке у два последња примера. То значи да је исти корак, 1000 опита међу векторима, у првом случају са већим отвором између „једнаких“ ($\vec{x}$ и $\vec{z}$) компензован дужим другим другим вектором ($\vec{x}$ и $\vec{y}$) а мањим углом између њих, у другом случају. Коефицијент нагиба линије регресије ($b = 0.005$) каже нам да нема значајне промене површине ($y = \Pi$) при промени броја опита ($x = M$).

За сада нема потребе да мењамо дефиницију дужина ових вектора. Та је опција иначе отворена не само на начин метрика претходно помињаних [19.], него и другачије. Важније је приметити да постоји необична аналогија између константности ове површине и Кеплеровог другог закона, на начин да се та површина не мења „удаљеношћу“, како променом броја опита M у овом случају, тако и дате планете од сунца у случају небеске механике. Да касније можемо покушати и остале аналогне формуле закона физике просто пренети у теорију вероватноће.

Пример 6. Расподела је из 3. примера, $\vec{p} = (0.6, 0.2, 0.1, 0.1)$, а почетно $M = 10\ 000$, са корацима $\Delta M = 1000$, као и дефиницију вектора $\vec{x}$ и $\vec{y}$ задржавамо. Други вектор је за корак испред првог.

M	$\vec{x}$	$\vec{y}$	Π
10 000	(5948, 2051, 984, 1017)	(6555, 2235, 1114, 1096)	43.41
11 000	(6650, 2132, 1102, 1116)	(7229, 2368, 1212, 1191)	52.82
12 000	(7079, 2443, 1278, 1200)	(7740, 2555, 1337, 1368)	131.40
13 000	(7742, 2634, 1293, 1331)	(8315, 2828, 1417, 1440)	28.60

Регресија ($y = a + bx$) има коефицијенте a = 24.77205 и b = 0.00342. ❑

Интересантно је приметити да „изјашњена“ расподела 5. примера, у односу на „неизјашњену“ из 4. примера, не мења ни површине ни њихове промене, бар не на олако видљив начин. Статистику од преко 10 000 узорака треба узимати скоро као извесност. Погледајмо још и сличан пример са осредњим изјашњавањем вероватноће.

Пример 7. Расподела је из 2. примера, $\vec{p} = (0.4, 0.3, 0.2, 0.1)$, а почетно $M = 10\ 000$, са корацима $\Delta M = 1000$, као и дефиниција вектора $\vec{x}$ и $\vec{y}$ непромењена, а други вектор је опет корак испред.

M	$\vec{x}$	$\vec{y}$	Π
10 000	(4010, 2983, 2033, 974)	(4419, 3309, 2210, 1062)	37.16
11 000	(4448, 3255, 2143, 1154)	(4731, 3612, 2424, 1233)	153.54
12 000	(4824, 3599, 2389, 1188)	(5299, 3850, 2553, 1298)	90.11
13 000	(5161, 3973, 2590, 1276)	(5619, 4198, 2781, 1402)	101.33

Регресија ($y = a + bx$) има коефицијенте a = -52.93640 и b = 0.01291. ❑

Откриће ове аналогије, површина које разапињу случајни вектори размака константног броја корака броја опита и површина два вектора планете од сунца размака константног интервала времена, даље сматрам (претпостављам) једнаким открићу гравитације. Овде открићу „силе неизвесности“, коју можемо видети у „тежњи“ природе да вероватније догађаје чини чешћима, да иде ка стањима мање емисије информације, мање комуникације, мањег дејства.

Можда то на први поглед делује сулудо, али не на други. Неизвесност је одбојна сила, свугде увек присутна, али веома слабог интензитета и на тај начин је разумљиво зашто је она током толиких векова развоја математике, наука и уопште филозофије остајала непримећена.

28. Центрифуга неизвесности

Центрифугалну силу (лат. „центар који бежи“) открили су и дефинисали Кристијан Хајгенс (1659) и Исак Њутн (1684). Тело се креће кружном путањом константне угаоне брзине ω на удаљености r, тангенцијалном брзином $v = r\omega$, кажемо ротира у некој равни око њеног центра O. Из Њутновог другог закона, да је сила једнака производу масе и убрзања тела, $F = ma$, налазимо $F = m\omega^2 r$ и то је центрифугална сила.

Да би површине које потег (вектор) од центра до тачке ротације у једнаким временима описивао увек једнаке површине, потребно је да брзина ротације опада са удаљеношћу, $vr =$ const, а две тачке у суседним интервалима времена, $d(\vec{x}, \vec{y})$, буду утолико ближе што је удаљеност већа. Тако разумемо да је производ rd приближно константан, као и поменута површина.

Додајмо овоме разматрања о централним кретањима (први део, Прилози), где сам доказивао да особина константне површине није привилегија гравитације (Кеплеров други закон, 1609), него је особина свих централних сила које се временом не мењају (извод по времену је нула), а затим и запажања претходног наслова овде [27.]. У претходном наслову, број понављања опита M биће оно што је удаљеност планете од сунца код Кеплера, а константно увећање броја опита ΔM биће константно време.

Константне централне силе, за које важи закон константне површине, нису само оне које опадају са квадратом удаљености (што се често, али погрешно подразумева). Али ако се такве смањују са квадратом удаљености, оне покрећу набој по коникама (елипсе, хиперболе, параболе) – докази су у другом делу о централним кретањима, моје скрипте „Прилози о теорији информације“. Додатно томе, један је став у физици, да константна централна сила опада са квадратом удаљености само ако се њено дејство шири константном брзином светлости.

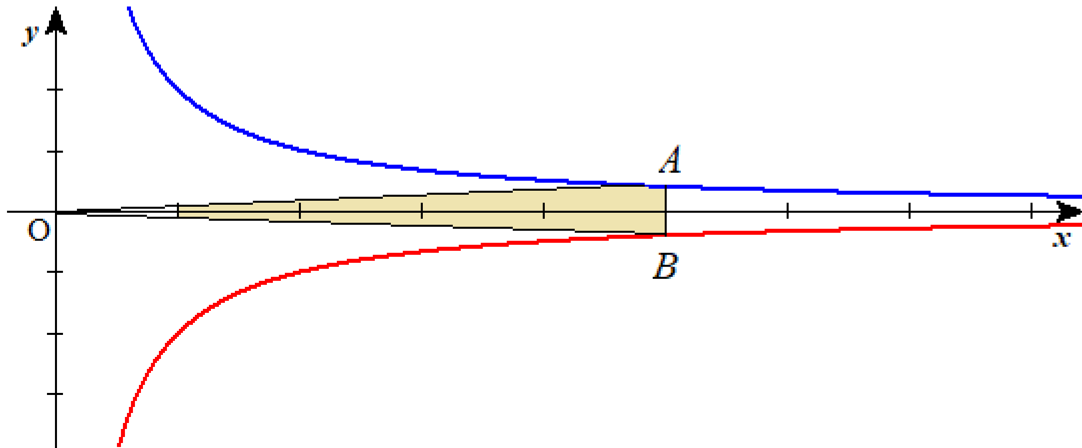
Да би разумели како се у то не уклапа ротација, замислимо једну велику платформу која се окреће око једне своје тачке, попут подијума рингишпила. Оно што ми замишљамо под ротацијом такве, је бесконачна раван чије све тачке ротирају истом угаоном брзином ω. Али јасно је да у стварности то није могуће, јер нису све тачке те платформе истовремене, нити су једнаковремене самој равни и било којем релативном посматрачу ван те ротирајуће равни. Бесконачно далеке тачке мировале би, а оне мало мање далеке каснилe би, заостајале у ротацији за ближима.

У исто се, на свој начин, не уклапа ни гравитација. Наиме, Меркур не ротира око Сунца по елипси. Тачније речено, његова замишљена „елипса“ обиласка полако ротира у правцу обритирања саме планете. Тако да ни гравитациона сила, прејака у близини Сунца, опада са квадратом удаљености, нити се гравитациони таласи тада крећу тачно „константном брзином“ светлости у вакууму. Иначе, једна секунда равна је путу од близу 300 000 km.

Посебан додатак истој аномалији Меркура, која се такође уклапа у Ајнштајнове једначине опште релативности, имам у теорији информације (Приче о информацији, 3. 30 Одложена гравитација) од пре неколико година. Простор памти и властита гравитација планете из њене ближе прошлости појачава се због мање брзине светлости и ротација убрзава. Синергијом перихел елипсе напредује брже него што би требао, у истом смеру камо и планета. Ефекат ће се појавити и другим избором четири просторно-временске координате из иначе 6-дим информатичког унивезума, а присутан је у релативистичким једначинама.

Ове аномалије помињем тек толико да нас не изненаде евентуалне особине „силе неизвесности". До тада покушајмо разумети зашто би она спадала у силе које набој покрећу по коникама. Нема ту пуно избора, пошто је сила о којој говоримо одбојна – коника њених набоја била би хипербола. А погледајмо и зашто она то заиста (приближно) јесте.

На слици су две хиперболе $y = \pm\frac{1}{x}$ у десној полуравни Декартовог правоуглог система координата (Oxy), горе плава са тачком A и доле црвена са тачком B, тако да троугао ABO одговара онима из претходних расподела. Угао $\sphericalangle(BOA) = 2\varphi$, приближно, где је угао $\varphi = \sphericalangle(\vec{z}, \vec{x})$, рецимо [26. Пример 1], са апсцисом као правцем очекивања $\vec{z}$, док је $\vec{x} = \overrightarrow{OA}$.



Троугао ABO састоји се од два правоугла троугла са заједничком катетом леђима окренутим, са укупном површином $\Pi = 2xy = \overline{OA} \cdot \overline{AB} \sin\varphi =$ const, јер је $y = \pm 1/x$. Дакле, путања којом би ишао вектор расподеле $\vec{x}$, претходних примера, хипербола је. Неком приликом ово такође можете тестирати генерисањем случајних бројева и цртањем, али водите рачуна да је трајекторија толико тачније хипербола ако тачнијом мером представљате дужине вектора. Последњи примери, какви су [27. Пример 1 и даље] биће довољни.

Сила мења вероватноће знамо од раније, а сада видимо да је и случајност сила. Оно што правило „константне површине" имплицитно наглашава, такође је да важи закон одржања неизвесности. Помињао сам то у више наврата, не колико закон одржања информације и сигурно не као друге такве законе , међутим овај нови није ништа мање битан. Без закона „одржања вероватноће" све би брзо клизило у извесност и остајало без информације (која је ткиво реалности и чија суштина је неизвесности), а сматрамо да се то не догађа.

Чаша је на столу, јер јој је такво стање највероватније, а биће тако следећих тренутака због закона одржања вероватноће, све док је рука или нека друга сила помери – реченица је коју сам ваљда у свакој књизи наводио, па и сада наглашавам – сила мења вероватноће, али и вероватноће су сила.

29. Закон одржања

Кеплеров други закон и поопштења казују нам о силама више него што смо се надали. Једно од нових сазнања, боље је рећи потврда већ претпостављеног, је закон одржања дејства (производа промене енергије и времена), а са њиме и одржања информације, те одржања неизвесности на начин како је меримо информацијом.

Када је у процесу константна повшина, $|\vec{x} \times \vec{y}| = xy \sin\varphi$ = const, јер одласком набоја од центра силе, повећањем дужина $x = |\vec{x}|$ и $y = |\vec{y}|$, смањујују се путеви $d(\vec{x}, \vec{y})$ које набој прелази током истих времена док је под утицајем даље силе. Смањивање путева набоја праћено је све мањим углом, чиме и синусом угла, а како оно прати промену дисперзије – одржавање исте површине повећањем броја понављања опита, M, има смисла посматрати као проблем синусне, односно угаоне дисперзије[6].

Повећањем дужина ће и скаларни производ, $\vec{x} \cdot \vec{y} = xy \cos\varphi$, расти јер му сва три фактора расту. Таквог га можемо разумети као акомулирану количину неизвесности. Начини израде вектора $\vec{x}$, $\vec{y}$ и $\vec{z}$ у примерима претходних наслова са све већим бројевима понављања опита, M, говоре нам о гомилању случајности и њеном чувању у компонентама вектора. Приметимо да тог чувања нема без закона одржања, па је „акомулација неизвесности“ у реченим векторима новост овог текста бар колико је и сам закон одржања.

Обе ново назване количине можемо чувати у облику комплексног броја

$$s = |\vec{x} \times \vec{y}| + i\vec{x} \cdot \vec{y} = xy(\cos\varphi + i\sin\varphi) = re^{i\varphi}$$

где за имагинарну јединицу важи $i^2 = -1$, а $r = xy$. У овој интерпретацији производ комплексних бројева, $s_1 s_2 = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}$, постаје запис нове акумулације $r = r_1 r_2$ и дисперзије $\varphi = \varphi_1 + \varphi_2$.

Пример нове интерпретације налазимо у старом запису слободне честице решења Шредингерове таласне једначине квантне механике $\psi(\vec{r}, t) = A\exp[i(\vec{k} \cdot \vec{r} - \omega t)]$. Овој A амплитуди таласа већој одговара већа шанса интеракције честице, $\vec{k}$ је таласни вектор, $\vec{r}$ текући положај честице, ω угаона фреквенција и t време. Сила неизвесности оспособљава квант дејством, додајемо.

Угао $\varphi = \vec{k} \cdot \vec{r} - \omega t$, упоређивањем са претходним, постаје стварна вредност кванта дејства. Он се такође може писати $\varphi = (\vec{p} \cdot \vec{r} - Et)/\hbar$, са импулсом $\vec{p}$, са енергијом E и Планковом редукованом константом $\hbar = 6.62607015 \times 10^{-34}$ џул секунди. Таласна функција писана у облику

$$|\psi(t)\rangle = Ae^{-\frac{iEt}{\hbar}}|\psi_E\rangle$$

одваја временски зависан фактор од решења временски независне Шредингерове једначине. Тој могућности третирања положаја и времена као посебних фактора треба придружити адитивност, која дозвољава сабирање више таласа-честица у једну

[6] Овом приликом измишљен назив.

$$|\psi(t)\rangle = \sum_j A_j e^{-iE_j/\hbar} |\psi_{E_j}\rangle$$

Ова позната сепарабилност и адитивност решења Шредингерове једначине, са интерпретацијама теорије информације, отвара могућност још једног начина разумевања тунелинга. О том начину заобилажења писао сам више пута детаљније (Solenoid), а третман је исувише нов и необичан да те описе сада скраћујем и банализујем.

У мноштву честица-таласа, захваљујући квантној суперпозицији и закону великих бројева настаје физичко тело. То је оно што видимо и из новог информатичког тумачења. Настаје „реалност" која је средња вредност мноштва имагинарности, неизвесности и занемаривања. Једнако како један крај нашег тела није у истој садашњости са другим (због ограничене брзине светлости и величине тела), тако и неки његови ситнији делови не долазе непосредно из „заједничке" прошлости, него нам стижу заобилажењем одлазећи из наше у псеудо-реалност (Dimensions) и враћајући се назад, или само долазећи однекуд отуда.

Живо биће за разлику од неживог има вишак информације. Оно се састоји и показаће више моћи и опција, него мртва твар данас познате физике која дословно следи принцип најмањег дејства. У комбинацији вишка слобода, закона одржања и тежњи ка њиховим редукцијама, интелигентно је понашање. Ми желимо спознати законе света око себе, желећи избећи „силу неизвесности" да би побегли у смирај и неделовање, вучени „инерцијом" која је заправо одраз начелног минимализма информације. Резултат су знања која откривамо, потреба за интелигенцијом, или цивилизовање у преносу истог преноса и гомилања (слободе) на колектив.

Покушаји потчињавања других себи, или себе другима, тражење је одушка од вишка слобода када то ослобађање од нас тражи начелни миминализам, а спречава закон одржања обзиром да је сва околна супстанца већ попуњена и жели исто што и ми. Ми тада личимо на олују којој нешто не да да се смири, упркос свеопштем принципу најмањег дејства којег следи сваки њен део, свака њена честица. Живот слично олуји има своју младост, зрелост и старост.

Сама интелигенција, пак, због истог (негде сам још читао да је, а ево зашто је) би да се испољава. Она тежи ситуацијама где би се могла исказати, местима више могућности. Интелигенција је тако ($I = S/H$) директно пропорционална обиму (количини) слобода својих интеракција са околином (S) и обрнуто пропорционална објективним ограничењима, хијерархији (H) околине. Према томе, обим интеракција ($S = I \cdot H$) директно је пропорционалан „интелигенцији" (способности ношења са опцијама околине) и „хијерархији" (способности околине да смањује опције субјекта). Отуда и информација перцепције, а и сложенија формула од ове када се исто примени у многострукости.

30. Комутатори

Елементарне особине комутатора биле су предмет (моје) теорије информације одавно, просто јер изражавају интензитет векторског производа вектора и површину коју разапињу два вектора. Када су вектори дужи, а њихов распон већи, као што смо видели, они могу симулирати Кеплеров други закон и рећи нам нешто о дејству, односно о информацији коју својом константном површином представљају.

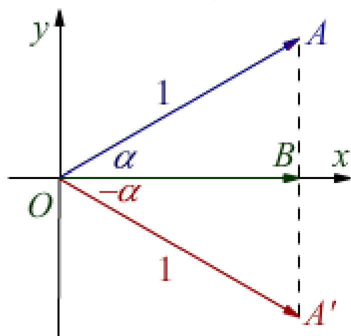
Захваљујући рачуну комутатора (Coordinates of inscribed circles in a triangle, February 13, 2015) могу се изражавати разне формуле, од елементарне геометрије преко алгебре па све до (нове) теорије информације. Срећом сам поједине записао у тим прилозима и тако сачувао. Неке ненадано могу наћи на интернету ових дана, али сећам се и других које још увек могу реконструисати из наставе и тадашњих разговора. Верујем да следећи од таквих примера још увек нема на другим местима.

Ма колико били дугачки низови координата, њихове n-торке бројева које припадају истом n-дим простору, два вектора разапињу један паралелограм и леже у једној равни. Нека су такви вектори $\vec{a} = (a_x, a_y)$ и $\vec{b} = (b_x, b_y)$, интензитета $a = |\vec{a}|$ и $b = |\vec{b}|$ и угла међу њима $\varphi = \sphericalangle(\vec{a}, \vec{b})$. Први тако је њихов скаларни производ, други интензитет њиховог векторског производа, односно комутатор $[\vec{a}, \vec{b}] = a_x b_y - a_y b_x$:

$$\begin{cases} \vec{a} \cdot \vec{b} = a_x b_x + a_y b_y = ab\cos\varphi \\ |\vec{a} \times \vec{b}| = a_x b_y - a_y b_x = ab\sin\varphi \end{cases}$$

Задали смо шему означавања коју даље користимо са различитим словима и бројевима.

Пример 1. Нека су два вектора јединичне дужине постављена симетрично око апсцисе, на слици лево, тако да је:



$$\overrightarrow{OA} = \vec{a} = (a_x, a_y), \quad \overrightarrow{OB} = \vec{b} = (b_x, 0),$$
$$\overrightarrow{OA'} = \overrightarrow{a'} = (a_x, -a_y), \quad a = a' = 1$$

и нека је $\vec{b}$ висина троугла $O\vec{a}\overrightarrow{a'}$. Углови које вектори a и a' граде са апсцисом (x-осом) су α и $-\alpha$. Тада је:

$$\frac{-a_x a_y - a_y a_x}{aa'} = 2 \cdot \frac{-a_y b_x}{ab_x} \cdot \frac{a_x b_x}{a' b_x}$$
$$\sin 2\alpha = 2\sin\alpha\cos\alpha$$

Доказали смо адициону формулу за синус двоструког угла. ❑

Пример 2. Са исте слике читамо даље:

$$a_x^2 - a_y^2 = \left(\frac{a_x b_x}{b_x}\right)^2 - \left(\frac{-a_y b_x}{b_x}\right)^2$$

$$\cos 2\alpha = \cos^2\alpha - \sin^2\alpha$$

Доказана је адициона формула за косинус двоструког угла. ❑

Општу адициону формулу за синус збира углова можемо добити помоћу било која три вектора $\vec{a}$, $\vec{b}$ и $\vec{c}$ у равни Oxy, таква да се углови између њих надовезују: $\alpha = \sphericalangle(\vec{a},\vec{b})$, $\beta = \sphericalangle(\vec{b},\vec{c})$. Додајмо да су дужине тих вектора јединичне, што неће умањити општост доказа.

Пример 3. Уврштавамо компоненте у адициону формулу за синус збира и израчунавамо:

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$$

$$\frac{c_x a_y - c_y a_x}{ac} = \frac{b_x a_y - b_y a_x}{ab}\cdot\frac{c_x b_x + c_y b_y}{cb} + \frac{b_x a_x + b_y a_y}{ab}\cdot\frac{c_x b_y - c_y b_x}{cb}$$

$$c_x a_y - c_y a_x = \left(a_y b_x^2 c_x + a_x b_x b_y c_y - a_x b_x b_y c_x - a_x b_y^2 c_y\right) + \left(a_x b_x b_y c_x - a_x b_x^2 c_y + a_y b_y^2 c_x - a_y b_x b_y c_y\right)$$

$$c_x a_y - c_y a_x = a_y\left(b_x^2 + b_y^2\right)c_x - a_x\left(b_x^2 + b_y^2\right)c_y$$

а обзиром да је $b = 1$ имамо тражену једнакост. ❑

Датим редоследом израчунавања пример потврђује тачност наведене опште адиционе формуле за синус збира углова, а обрнутим редоследом доказали бисмо је, јер наведене једнакости међу собом су еквивалентне (када је тачна једна тачна је и друга и обрнуто). Евентуалном читаоцу овог текста остављам да аналогно, ради вежбе, изведе општу адициону формулу за косинус зибра или разлике углова.

Линеарни оператори над векторима такође чине „векторски простор“. Он је кажемо дуалан датим векторима и, наравно, са одговарајућим скаларним производима вектора и комутаторима. Књига Квантна механика која је била одобрена за штампу почетком 2018. године, али сам одустао јер „ко би то читао“, садржи и класичне доказе особина оператора ротације (Пример 1.1.62) и понеке још занимљивости које такви уџбеници обично немају. Међутим, мој доказ (помоћу комутатора) наћи ћете у књизи Дејство информације, Економски институт Бања Лука, 2021. године (2.5.5 Ротације).

Пример 4. Вектор $\vec{a} = \left(a_x, a_y\right)$ ротиран је у вектор $\vec{b} = \left(b_x, b_y\right)$ за угао φ, око почетка O правоуглог Декартовог система координата (Oxy). То значи да су дужине ова два вектора једнаке ($a = b$), али оне не морају бити јединичне. Знамо да је:

$$\begin{cases} ab\cos\varphi = a_x b_x + a_y b_y \\ ab\sin\varphi = a_x b_y - a_y b_x \end{cases}$$

Множимо прву једначину са a_x другу са $-a_y$ и саберимо их:

$$aba_x\cos\varphi - aba_y\sin\varphi = a_x\left(a_x b_x + a_y b_y\right) - a_y\left(a_x b_y - a_y b_x\right) =$$

$$= a_x^2 b_x + a_x a_y b_y - a_x a_y b_y + a_y^2 b_x = a_y^2 b_x + a_x^2 b_x$$

$$= \left(a_y^2 + a_x^2\right)b_x = a^2 b_x = abb_x$$

Скраћивањем са ab налазимо једнакост $a_x\cos\varphi - a_y\sin\varphi = b_x$. Множењем прве горње са a_y а друге са a_x налазимо:

$$aba_y\cos\varphi + aba_x\sin\varphi = a_y\left(a_x b_x + a_y b_y\right) + a_x\left(a_x b_y - a_y b_x\right) =$$

$$= a_x a_y b_x + a_y^2 b_y + a_x^2 b_y - a_x a_y b_y = a_y^2 b_y + a_x^2 b_y$$

$$= \left(a_y^2 + a_x^2\right) b_y = a^2 b_x = abb_y$$

Па скраћивањем са ab, опет, сада налазимо $a_x \sin\varphi + a_y \cos\varphi = b_y$. Укратко, две ове једнакости можемо писати матрично

$$\begin{pmatrix} b_x \\ b_y \end{pmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{pmatrix} a_x \\ a_y \end{pmatrix}$$

А то је тражена трансформације вектора $\vec{a}$ у $\vec{b}$ ротацијом. ❑

За разлику од ових „тривијалности“, комутатори се чешће и озбиљније користе у квантној физици, прво за Хајзенбергове релације неодређености. Ево једног шире познатог таквог примера.

Пример 5. Држећи се правца апсцисе, оператори положаја и импулса квантне честице су просто множење таласног пакета ψ положајем ($\tilde{x}\psi = x\psi$) и узимање извода ($\tilde{p}\psi = -i\hbar\frac{\partial\psi}{\partial x}$). Комутатор је:

$$[\tilde{x}, \tilde{p}]\psi = \left(x\frac{-i\hbar\partial}{\partial x} - \frac{-i\hbar\partial}{\partial x}x\right)\psi = -i\hbar x\frac{\partial\psi}{\partial x} + i\hbar\frac{\partial}{\partial x}(x\psi) =$$

$$= -i\hbar x\frac{\partial\psi}{\partial x} + i\hbar\psi + i\hbar x\frac{\partial\psi}{\partial x} = i\hbar\psi$$

Како то важи за сваку таласну функцију ψ, биће $[\tilde{x}, \tilde{p}] = i\hbar$, а отуда даље принцип неодређености, да из некомутативности следи већа неодређеност положаја (веће је Δx) када имамо импулс боље одређен (мање Δp) и обрнуто, при чему важи неједнакост $\Delta x\, \Delta p \geq \hbar/2$. ❑

Још сличних примера налази се у мојој Кантној Механици, или другде. Неизвесност (информација) је у посебној спрези у случају импулса и положаја (енергије и времена). Смањивањем једног расте друго и обрнуто у повезаности која, према теорији информације, информацију и физичко дејство чини еквивалентнима.

31. Ротације

Ако физику разумемо као науку која проучава закона одржања, што није далеко од истине, онда су изометрије њене главне алатке из математике. То су пресликавања чувају удаљености параова тачака, попут рефлексија, транслација и ротација. Физичари их воле називати симетријама, такође их све можемо називати ротацијама.

Наиме, централна симетрија је ротација за $180°$ око „центра“. Осна симетрија је ротација за $180°$ око „осе“. Огледалска рефлексија је ротација (у додатној, замишљеној димензији простора) око равни огледала. Транслација се састоји од две централне симетрије, а ротација је ротација. Идеја која из тога произилази је значај квантна дејства [30. Пример 5] у формализму ротацијама.

Када бисмо под „ротацијама“ подразумевали само реалне ротације [30. Пример 4], игнорисали би додатне временске димензије које су важан део ове теорије информације (Complex). Комплексни домен припада псеудо-реалности, у коју и из које микро-реалности знају бежати (Bypass), без које би „информатичка теорија“ била каузална, што она није. Према томе, главне ротације на које сада мислим су множења комплексним бројевима.

Силе неизвесности [27.] покрећу догађаје, попут олује су, или мешања кафе, млека и шећера да би у шољи добили „добро“ стање. У процесу смиривања су „позитивна“ стања, која препознајемо, на која смо навикли временом живећи са принципијелним минимализмом информације. Она постају пожељна за економски и неке друге развоје друштава, јер са другачијима не умемо. Те корисне и стабилне ситуације попут су лепог времена које долази после олује, којег не би било без олуја. Као поплаве реке Нила Древног Египта оне разарају су па стварају.

Економски просперитети (као и све, али баш све природне појаве) долазе у таласима, а друштва која знају некако толерисати разорне периоде тих таласања, успевају доживети њихове највеће благодети. Уосталом, зато друштва која почну да се превише уређују праве свој мртвачки сандук (Democracy). Зато ће једнако вешт играч тактике „жртвом до победе“ бити успешнији од равног себи који би само спајао „добро са добрим“ у компромисима (Win Lose).

Све ове појаве, наизглед неповезане, последица су „ротација“ на микро нивоу физике, које уносе неизвесност и покрећу овај свет. Оно што бисмо требали је приметити их у комутатору положаја и импулса квантне честице, управо у износу $i\hbar$ који би могао бити њихов минимум. Ако је тако, тај је износ раван синусу минималног угла неизвесности евентуално помножен неком константом. Сама ротација чији је тај квант део, $e^{\pm i\varphi}$, потенцијални је фактор сабирака информације перцепције.

Задржавам се на објашњењима којих ротацијама недостаје више од рачунских метода. Оне овако могу изгледати очигледне колико онако недохватљиве. Претпоставка је да постоје ротације веома малог квантног угла φ (приближног синусу), реда величине $\hbar = 6.626176 \times 10^{-34}$ Js са наизглед незнатним утицајем на релативно огромну реалност. У складу је то са законима великих бројева, а и са интуитивним одбијањем као и постојање атома чије чулне особине нам нису очигледне.

Имагинарне дисперзије 10^{34} пута су мање од џул секунде и могу нам изгледати нестварне, давати нам илузију „чисте реалности“. А тако су мале због „сила неизвесности“, претпостављам. Реалност се понаша као мехур од сапунице у окружењу псеудо реалности. На окупу начелним минималмом информације, као капиларним силама и веома слабим али свеприсутним и упорним, реалност се

држи по страни, или рецимо опире имагинарности управо снагом произведеном силом, односно агресијом неизвесности.

Према томе, информацију перцепције могли бисмо дефинисати и као спрегу вектора који садрже имагинарне ротације, иначе занемарљиво малих неизвесности α_k и β_k, за $k = 1,2,\dots,n$, тако да је $\vec{x} = \left(\xi_1 e^{i\alpha_1}, \xi_2 e^{i\alpha_2}, \dots, \xi_n e^{i\alpha_n}\right)$ и $\vec{y} = \left(\eta_1 e^{i\beta_1}, \eta_2 e^{i\beta_2}, \dots, \eta_n e^{i\beta_n}\right)$. Тада би информација перцепције, у општијем случају, била писана у облику:

$$S = \langle \vec{x} | \vec{y} \rangle = \vec{x} \cdot \vec{y}^* = \sum_{k=1}^{n} \xi_k \eta_k e^{i(\alpha_k - \beta_k)} = \rho e^{i\varphi}$$

где опет дозвољавамо могућност (углавном занемарљиве) имагинарности, $\varphi \neq 0$. Подразумева се да су бројеви редом α_k и β_k, као и број ρ реални, а такође да многи од углова могу бити нуле.

Међутим, ова интерпретација информације перцепције веома је квантно-механичка, доследна би се превише односила само на вероватноће. Наиме, резултат $\rho e^{i\varphi} = \rho \cos\varphi + i \sin\varphi \to \rho \cos\varphi$, јер $\varphi \to 0$. Први сабирак тада говори о шансама догађања ове спреге, а други о њеној информацији и количини неизвесности. То је, наравно, према самој природи овог тумачења које није једино.

Видели смо да већ из такве интерпретације, вероватноће S, број $I = -\log S$ је информација. Али и у било којој другој интерпретацији, када би непосредно I била информација спреге, добијали би и $S = e^{-I}$ као вероватноћу. У оба случаја, тумачењима резултата требала би накнадна подешавања, јер у крајњем случају (микро света), не може се побећи од имагинарности изворне неизвесности и периодичности резултата које доноси употреба комплексних бројева.

Чињеница да таласну природу материје већ имамо у квантној механици, на какав такав примећен начин, Де Бројевом хипотезом (1924) и Шредингеровом таласном једначином (1926), разумевање ових идеја теорије информације чини лакшим послом.

III Део: Континуум

Банална наука била би она која се ослања само на непосредне чулне опажаје. То са становишта теорије информације која полази од слојевитости „неизвесности“, такође од њене релативности и објективности. Објаснићу.

Слојевитост „непознатог“ појављује се већ у ситуацијама када неки субјекат нешто не зна и то се промени након обавештења. Релативност је, на пример, појава различитих фреквенција светлости зависно од брзине њеног извора у односу на посматрача. Објективност „незнања“ је немогућност сазнања у датим околностима. Поред тога, ова теорија полази од претпоставке да је информација тканиво простора, времена и материје, а неизвесност њена суштина, тако да је „обавештеност“ свега о свему на неки за сада науци несхватљив начин, рецимо неизвестан, ипак могућа.

Окосница те повезаности је „истина“. Оно што бисмо могли доказати да се не може десити – неће се десити. Оно што бисмо доказали да је тачно – не можемо променити. Сваки део (математике), свака област тачних сазнања у сагласности је са сваком другом тачном облашћу и полазећи од ове „необичне“ особине истина можемо говорити о једном „свету истина“, насупрот „света неистина“. При томе, та два еквивалентна су (в. Dualism of lies), а лаж је потиснута истина, слабија због својег мање видљивог „ткива“, али привлачнија због спонтане тежње ка мањој информацији, или што је исто ка већој вероватноћи (The Truth).

Према реченом, истине су објективније од онога што сматрамо „физичком реалношћу“. У то нас наука уверава стално, кроз векове, поправљајући се и откривајући на нове и нове начине. Оно што се у тим процесима примицања истини не мења, него се само допуњују, математичке су теореме. А сада додајмо и „објективност“ у смислу немогућности мењања.

Објективност неизвесности можете покушати разумети помоћу Геделове теореме некомплетности (Incompleteness), да сваки математички систем има њему недоказивих а иначе истинитих исказа. А једноставније, приметно је то у опису бесконачности скупа природних бројева $\mathbb{N} = \{1, 2, 3, \dots\}$, када подразумевамо да од сваког унапред датог природног броја има већи, у Раселовом парадоксу, да нема скупа свих скупова (од сваког „универзалног скупа“ има већи).

Овај начин, разумевања неизвесности о којој говорим, битан је за поентирање које следи, па даље настављам набрајати са дефиницијом непрекидности функција, или рецимо конвергенције низа бројева. Кажемо да низ a_n конвергира броју a, ако за свако $\varepsilon > 0$ постоји $n_0 \in \mathbb{N}$ такво да сваки $n \in \mathbb{N}$ и $n > n_0$ повлачи $|a_n - a| < \varepsilon$. Ако је овакав услов конвергенције испуњен, писаћемо или $\lim_{n\to\infty} a_n = a$, или $a_n \to a$ када $n \to \infty$, односно рећи да је a гранична вредност (лимес) низа a_n.

У сваком од набројаних случајева приметимо да „од унапред датог има даље“. Затим то опазимо и у тврђењу, рецимо, да за сваку дату енергију тела постоји становиште, или релативни посматрач који ће је видети већом. Извор светлости, фотона енергије $E = hf$, где је $f = 1/\tau$ фреквенција, а τ је трајање једне осцилације, који се приближава релативном посматрачу, имаће увећану енергију, односно фреквенцију, што значи краћа трајања осцилација (Doppler effect) и то скраћивање може ићи у недоглед. Слично скраћивање налазимо посматрајући таласне дужине долазећих таласа. У том смислу су простор и време бесконачно дељиви.

Следећи корак, за почетак, много је тежи јер се тиче наших предрасуда. Еволуција нас је опремила чулима. Она су наши прозори у свет и средства преживљавања у разним окружењима. Нама могу изгледати довољна за опстанак, а онда нам је и лако поверовати да нам сама чула дају представу о свему. Тако настају идеје да физика треба бити независна, рецимо од логике, или математике, те да се треба ослањати само на мерења. Штавише, да из физике треба избацити све друго осим непосредно мерљивог (Ернст Мах, 1838—1916, немачки физичар и филозоф).

Када је Лудвиг Болцман (1844—1906) почео заговарати титрање молекула у преносу топлоте и температуре на хладнију околину, на једном његовом предавању Мах је устао и окренут публици узвикнуо: „Немојте слушати овог човека, овај човек је будала, атоми не постоје!" Мах је заговарао филозофију према којој физички није реално оно што не можемо видети, чути, опипати, а атоми су тако малени да то никада неће бити могуће. Знамо да је физика у наставку „опипала" атоме, али и да се још увек држи „чулне реалности", додуше сада на дугом штапу.

Верујем, већ препознајете да овим лаким уводом најављујем реалност, или бар псеудо-реалност бесконачности. Ако је тако у праву сте, али при томе не мислим само на бесконачност природних бројева, алеф нула ($\aleph_0$), на тзв. пребројиву бесконачност, већ и на непребројиву (континуум) скупа реалних бројева, за коју су оснивачи теорије скупова (пре свих Георг Кантор) доказали да је већа од пребројиве. Такође и на остале бесконачности, али оне неће бити тема овде.

Међутим, са становишта теорије информације реалност је попут неизвесности, она је такође, али и на свој начин слојевита, релативна и објективна. Стални ток садашњости, из будућности појединца у његову прошлост, врста је слојевитости. О постојању сличне врсте „псеудо реалности" потврде су и додатне димензије времена (Dimensions). Оба примера говоре и о релативности реалности, зато што тела немају јединствену садашњост, због своје величине и ограничене брзине светлости. Ова релативизирања објективности различите реалности третира једнако објективнима.

Разлике између „ове и оне" реалности тако праве комуникације, као и закон одржања. Прво нам је јасно, надам се, а ово друго долази из основне разлике између бесконачности и коначности. То је особина бесконачности да њен прави део буде једнак целини, који коначни скупови немају. Из тога следе закони одржања, теорема Еме Нетер, симетрије и принцип најмањег дејства.

Са друге стране, принцип најмањег дејства еквивалентан је начелном минимализму информације, да природа спонтано тежи мање информативним системима, а оба еквивалента су принципијелно већој вероватноћи (чешћем реализовању вероватнијих исхода). Тако долазимо до становишта да су „реалности" острва коначности у бесконачности. Оно што свет наших перцепција држи даље од тих бескрајних неизвесности управо су „закони коначности", дакле закони одржања, инерција, ток ка мањој комуникацији.

32. Бесконачност

Универзалност и неизвесност информације рећи ће да ма како велика била (замишљена) васиона, она је такође нека информација. Свеукупност у коју бисмо поверовали да постоји, или само хтели да замислимо, опет је јединствена и отворена како према унутрашњем тако и нечем даљем. Она је немогућа као крај свега, а могућност покушаја замишљања ње само је још једна потврда лажи као прикривене, кодиране истине (The Truth).

То „откриће“ ове теорије информације заправо и није новост, бар не у математици. Мислим на Геделове теореме о непотпуности (1931), да не постоји потпун и конзистентан формални систем који коректно описује природне бројеве и да ниједан довољно строг систем који их описује неће моћи потврдити своју сопствену конзистентност. При томе, у математичкој логици, формални се систем сматра конзистентним ако не садржи контрадикције (за сваку пропозицију, не могу бити доказиве и она и њена негација), а систем је потпун који је довољан за изградњу целе тражене теорије.

Закон одржања информације, пак, следи из коначности (наших) опажаја. Наиме, ако би (нам) било могуће бесконачно делење информације, онда би и њена већа подела била еквивалентна мањој и не би могао важити закон одржања. Зато је информација коју опажамо у пакетима (Packages), ми тога били свесни или не. Теореме и ставови математике су дискретни скупови, кораци доказа, као и правне ставке. Такође, кванти физичког дејства који су еквивалентни информацијама и најмањи носиоци тих пакета (информација је количина), увек су неки пакети.

Сада прво прочитајте о Хартлијевој информацији [05.], па настављамо. Ту је идеја о могућностима проширења зависности између броја једнако вероватних могућности $N = 1, 2, 3, \dots$ и информације $I = \log N$ коју они носе. Важне идеје Георга Кантора у математици су скупови и бесконачности. Да бесконачне величине, назване кардиналним бројевима, треба упоредити бијекцијама (обострано једнозначним пресликавањима), својевремено је било много теже прихватити од самих скупова. Након „Раселовог парадокса“, да не постоји „скуп свих скупова“, који је претио да уруши концепт саме теорије, примећено је да „од сваке уније скупова има већа“ и да се са тим постулатом може ићи даље, остала је привремена пометња око Канторовог упоређивања величина бесконачности као и то да њихових величина има разних. Ако постоји бијекција између елемената два скупа, тада и само тада, Кантор је коначно утврдио да скупови имају једнаке (кардиналне) бројеве елемената.

Детаљније о упоређивању бесконачности можете наћи и у мојој (лакшој) књизи „Многострукости“ (2018) у наслову „Различитости“ на странама 11-14. Покушајмо сада те две (Хартлија и Кантора) идеје удружити. У логаритму базе 2, јединице „бит“ информације, број могућности је $N = 2^I$. То значи да ће у низу $I = 1, 2, 3, \dots$ бинарних цифара писаних иза нуле и запете $0, b_1 b_2 \dots b_I$, где свако $b_k \in \{0, 1\}$ за свако $k = 1, 2, \dots, I$, бити тачно $N = 2^I$ могућности за записе различитих позитивних бинарних бројева мањих од један. Наиме, на свакој од I позиција постоје две могућности, а два пута два ... I пута једнако је 2^I.

Доследно даље, када за информацију узмемо кардинални број пребројиво бесконачног скупа $I = \aleph_0$ (алеф нула), број могућности биће континуум. Можемо рећи и овако: ако у претходном, коначном резултату, пустимо да $I \to \aleph_0$ онда $N \to$ континуум. Просто тако.

Ово „једноставно“ удруживање идеја Хартлија и Кантора поопштавање је сваке од њих и сасвим је нови концепт ако га наставимо интерпретирати у складу са (мојом) теоријом информације. На два начина можемо посматрати бројеве реализованих могућности: са становишта само једног субјекта и са становишта свих постојећих.

У претходном објашњавању методе инфинитезималног рачуна, да „од унапред датог има даље“, која омогућава какву-такву непрекидност простора и времена (о детаљима другом приликом), до пребројиве бесконачности простор-времена мали је корак. Толико што се тиче уситњавања, макар до дискретне мреже положаја и тренутака попут система координата самих рационалних бројева.

Наиме, између свака два рационална броја ($a, b \in \mathbb{Q}$) постоји рационални број ($\frac{a+b}{2} \in \mathbb{Q}$), кажемо скуп рационалних бројева је „свуда густ“, али низ формиран од све више коначно узетих децимала броја $\pi = 3{,}14159\ldots$ састојаће се само од рационалних бројева и конвергираће броју π који није рационалан. За разлику од пребројиво бесконачног скупа рационалних бројева $\mathbb{Q}$, пример скупа чији сви низови конвергирају унутар самог скупа био би континуум реалних бројева $\mathbb{R}$.

Са друге стране, простирање простора ка даљинама и трајање времена, бесконачни су. Узимајући становишта различитих посматрача долазимо до бесконачне дељивости простора и времена, као код рационалних бројева, ако и само ако имамо бесконачно простирање и трајање, или довољно могућих посматрача. Ових других мора бити „довољно“ за сваког од субјеката дате реалности.

Поновићу још једном да „реалност“ у теорији информације третирам помоћу комуникације. Ако А може непосредно комуницирати са Б, онда су А и Б непосредно узајамно реални. Ако А и Б нису у могућности да комуницирају непосредно, али постоји Ц са којим то оба од њих могу, онда су А и Б посредно узајамно реални. Ово је опис мреже (чворова и повезница) за дефинисање реалности.

Са изложеним поставкама долазимо до открића да је реалност реализивана информација околних неизвесности. Исходи пребројиво бесконачно могућности „унутар су“ околног континуума. Опет то је овде понављање, ако се сетимо сила неизвесности [27.], наше реалности коју окупља принцип минимализма – насупрот одбојних вањских неизвесности. Тежња мањој информацији упоредива је, не превише, са привлачним капиларним силама због којих се мехур од сапунице не распада.

Да је наша реалност реализивана информација околних неизвесности видимо на још један начин. Посматрајмо низ могућности исхода, на које наилазимо током развоја садашњости, као позиције децималног записа $0, d_1 d_2 d_3 \ldots$ позитивног реалног броја мањег од један. Без обзира што је број позиција пребројиво бесконачан, број могућих записа је континуум. Број позиција еквивалент је броју елемената наше реалности, а број могућности остатак је 6-дим простор-времена.

33. Амплитуда и таласна дужина

Амплитуда, таласна дужина и брзина формално дефинишу талас. За разлику од продирања таласа кроз средине различитом брзином (Speed of light), овде је акценат на прве две карактеристике. Та брзина за светлост је приближно $c = 300\ 000$ km/s, нормална брзина звука $c = 340$ m/s, а рецимо брзина воденог таласа метра дубине је $c = 3$ m/s. Каква год, сматраћемо брзину непромењљивом, односно талас ћемо посматрати у средини где се она не мења.

За време t талас брзином c, крећући се напред или назад дуж истог прваца, пређе пут $x = \pm ct$. У позитивном смеру узевши биће стално $x - ct = 0$, у негатативном $x + ct = 0$. Другим речима, ако су f_1 и f_2 произвољне функције, тада је $f(x,t) = f_1(x - ct) + f_2(x + ct)$ константно, када је реч о таласима. Према томе, ако узмемо прве и друге парцијалне изводе, налазимо:

$$\frac{\partial}{\partial x}f(x,t) = \frac{\partial}{\partial x}f_1(x - ct) + \frac{\partial}{\partial x}f_2(x + ct) = f'_1(x - ct) + f'_2(x + ct)$$

$$\frac{\partial}{\partial t}f(x,t) = \frac{\partial}{\partial t}f_1(x - ct) + \frac{\partial}{\partial t}f_2(x + ct) = -cf'_1(x - ct) + cf'_2(x + ct)$$

$$\frac{\partial^2}{\partial x^2}f(x,t) = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}f(x,t)\right] = \frac{\partial}{\partial x}\left[f'_1(x - ct) + f'_2(x + ct)\right] = f''_1(x - ct) + f''_2(x + ct)$$

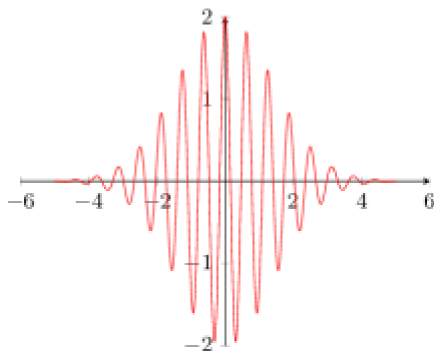
$$\frac{\partial^2}{\partial t^2}f(x,t) = \frac{\partial}{\partial t}\left[\frac{\partial}{\partial t}f(x,t)\right] = \frac{\partial}{\partial t}\left[-cf'_1(x - ct) + cf'_2(x + ct)\right] = c^2\left[f''_1(x - ct) + f''_2(x + ct)\right]$$

па можемо формирати парцијалну линеарну диференцијалну једначину другог реда

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = 0$$

То је таласна једначина дуж дате x-осе (апсцисе).

Таласна функција $f(x,t)$ коју смо управо дефинисали представља „интензитет“ који се креће са



таласом сталне брзине и који је зато константан. Можемо га сматрати произвољним делом фазе таласа, који сам по себи ништа не говори о амплитуди (највећем одступању од осе кретања).

Међутим, где је вредност те функције већа, талас се јаче опажа. Зато, а нарочито у случају микро-честица, можемо амплитуду одређивати и као вероватноћу обзервабле. На слици лево види се да је таква једна вероватноћа заправо густина вероватноће размазана од -4 до +4 апсцисе. То се управо и ради у квантној механици. Сличну датој, Шредингерову таласну једначину, детаљно сам решавао у „Квантној механици“, током многих делова књиге.

Решења таласне једначине задовољавају аксиоме векторског простора, што значи да су и она неки векторски простор. Таласне функције су вектори који према осама координатног система бираним да буду обзервабле затварају углове чији је збир квадрата косинуса јединица. Другим речима, ови квадрати косинуса представљају вероватноће неке расподеле. У векторским просторима скалара

комплексних бројева, каквим представљамо квантне системе, ови косинуси одговарају модулима амплитуда. То сам у поменутој књизи такође описао након Варадарајановог објашњења Борновог закона, да квадрат модула амплитуде представља вероватноћу мерења дате обзервабле.

Улога косинуса углова између вектора (таласне једначине) и координатних оса (обзервабли) може се разумети и као производ пројекције јединичног датог вектора на осу са пројекцијом јединичног вектра осе на дати вектор. Другим речима, то је скаларни производ нормираног датог вектора са ортовима (јединичним векторима координатних оса), односно врста „информације перцепције“, рецимо као перцепција самог себе у интерпретацији вероватноћа.

Другу врсту вероватноћа у вези са таласима дају таласне дужине. Гледајући концентричне кругове ширења таласа на површини воде видимо како им амплитуде опадају, а таласне дужине остају. То пре свега значи да су те две величине одвојене појаве. Прва одређује „видљивост“ и вероватноће опажања таласа, а друга њихову „размазаност“ на месту опажања.

У истој поменутој књизи квантне механике детаљно сам извео Комптонов ефекат, уз примедбу да он не доказује само честичну природу светлости, већ и бирање вероватније путање са становишта фотона. Обзиром да расипањем γ-зраке, након судара са угљениковим електронима, прелазе на путање већих таласних дужина, дакле већих размазаности и мањих вероватноћа на датом месту, то таласну дужину можемо сматрати показатељем вероватноће трајекторије. Већим се таласним дужинама придружују мање густине вероватноћа положаја фотона.

Та је књига формула квантне механике[7] са рецензијама завршена до априла 2018. године и имала је издавача и спонзора, никада није отишла у штампу, јер сам сматрао да је нема ко читати. Важна ми је у пдф-формату, на интернету, као збирка доказа, али и сведочанство о хипотезама каква је и ова о таласним дужинама. Исто тумачење таласних дужина, уз можда још интересантније обраде, можете наћи и у ранијим мојим радовима (Простор-Време, Комптонов ефекат), али га наводим и овде, јер се временом показује све бољом и још увек је, чини ми се, непознато у физици.

Наравно, изворне Хајзенбергове релације неодређености (у истим књигама) имплицитно користе ово својство таласних дужина. Положај честице размазан је по „дужини осцилације“, као и смеру и интензитету импулса, тако да смањујући прву величину добијамо тачније одређење положаја честице-таласа, али утолико нетачније меримо њен импулс и обрнуто, прецизирајући импулс у мерењу губимо на прецизности у одређивању њеног положаја.

[7] Р. Вуковић: Квантна Механика – репрезентација континуума, метричког и векторског простора (2018).

34. Бинарне расподеле

Када бацимо фер новчић вероватноће да падне писмо или глава једнаке су, $p = q = 1/2$. Али и ако новчић није „фер“, ове вероватноће су позитивни бројеви јединичног збира, па је само једно бацање (следећа) расподела вероватноће.

Пример 1. Проста бинарна расподела вероватноће:

$$F_1 = \begin{pmatrix} A & B \\ p & q \end{pmatrix}$$

Вероватноће случајних догађаја су $P(A) = p$ и $P(B) = q$, где су $p, q \geq 0$ и $p + q = 1$. ❑

Мало мање једноставни су случајеви понављања ове „просте бинарне расподеле“ са немењаним условима на различите начине. На пример, бацајмо новчић све док не падне „писмо“. Ако је тек у n-том покушају десио исход A, онда се у претходних $n - 1$ десило B, па је вероватноћа баш таквог догађај $p_n = q^{n-1}p$. Они су дисјунктни и збир им је један ($p_1 + p_2 + p_3 + \cdots = 1$), што значи да је ово још једна расподела вероватноћа.

Пример 2. Понављање до првог жељеног исхода

$$F_2 = \begin{pmatrix} A_1 \; A_2 \cdots A_n \cdots \\ p_1 \; p_2 \cdots p_n \cdots \end{pmatrix}$$

где је A_n догађај да ће се у n-том понављању претходног опита (F_2) по први пут десити исход A. Из тога следи вероватноћа $p_n = q^{n-1}p$, са јединичним бесконачним збиром. ❑

Средње вредности, варијансе и информације ових расподела израчунавао сам у књизи „Физичка Информација“ и нема потребе да та извођења сада понављам. Оно што би могло бити од значаја, а тамо није помињано, Био би следећи формални задатак.

Пример 3. Са вероватноћом p стање A прелази у стање B, а обрнуто, са вероватноћом q стање B у A. Поставља се питање вероватноћа прелазака $A \to B \to A$, затварања једног циклуса прелазака. Да појендоставимо задатак, ставимо $q = 1 - p$, па први прелазак $A \to B$ очигледно постаје случај F_2, дакле нешто што ће се сигурно кад-тад десити, али такође и други прелазак. Ма како задавали вероватноће p и q увек ће (до након бесконачно корака) A постати B које ће се потом вратити у стање A. ❑

Овај пример је интересантан у случају да се тако фермиони претварају у бозоне, односно честице супстанце у честице простора. Негде у међувремену приметили бисмо да има све мање супстанце, или бар да је она све више разређена у све већем простору. Мање честица супстанце не повлачи и мању њихову укупну енергију. Међутим, за разлику од последњег примера једне једине честице, поврат свих у тачно првом циклусу није известан догађај (Growing).

Математика је апстрактна, она издваја форму која би се дала мултиплицирати у, на друге начине, неповезане појаве што је са једне стране чини веома „суштинском“, а са друге „неузрочном“. Знам да се очекује управо обрнута употреба ових речи, да је математика узрочна и да није суштина, али размислимо! Ако се нешто односи на толико много неповезаних других ствари, на пуно наставака, онда оно није добар „узрок“ за предвиђање. Такође, оно је „основа“ многих ствари.

У овом смислу „суштинска природа“ математике требала би нам рећи још много новога о „добро познатом“ закону великих бројева [26. Дисперзије]. На пример, развојем из микро света у макро, физички систем постаје све извеснији (како се број константних случајних опита понавља, њихова средина узорка, просек, приближава се њиховој теоријској средини). Било би да се информација датог повећаваног система смањује да није умножавања других околности, каже овај Бернулијев закон (1713), који можемо потражити и у самом простору и времену.

Знамо да се памћење прошлости налази у материјалним траговима садашњости. На основу закона појединих (форензичких) наука откривамо шта је било и реконструишемо догађај кога више нема. Оно што не знамо је да би се још неки од шаблона из тога могли извући и генерализовати, па ево предлога за њих неколико.

Прво питање, пошто су у макро свету исходи случајних догађаја извеснији, може ли се доследно и начелна тежња природе ка већој вероватноћи, или мањој информацији, или комуникацији, а и дејству, на неки начин појавити у виду „повећања“ овог света? Друго питање, може ли се слична појава затим јављати и као процес, у току времена из наше прошлости ка будућности? Одговор би могао бити у „спонтаном“ ширењу свемира (онолико колико је спонтана тежња физичких система ка мање информативним), те у одговарајућем току садашњости ка већој извесности.

То су необавезне претпоставке које тек треба расправљати, што ћемо и радити али једном другом приликом, а за сада прелистајмо још неколико важних познатих ставова бинарних расподела. Нај познатија бинарна расподела свакако је одговор на питање, у серији од $n = 1, 2, 3, \dots$ понављања случајног догађаја из примера 1, колика је вероватноћа да се A деси тачно $k = 0, 1, \dots, n$ пута?

Пример 4. Понављамо F_1 (пример 1) n пута. Вероватноћа да се A деси сваки пут је p^n, а ниједном је q^n. У низу од n могућих, k исхода може се остварити на $\binom{n}{k}$ начина, а свака од тих комбинација састоји се од k реализованих и $n - k$ нереализованих догађаја A. Према томе, вероватноћа да се у низу од од n покушаја оствари њих тачно k износи $P_k = \binom{n}{k} p^k q^{n-k}$.

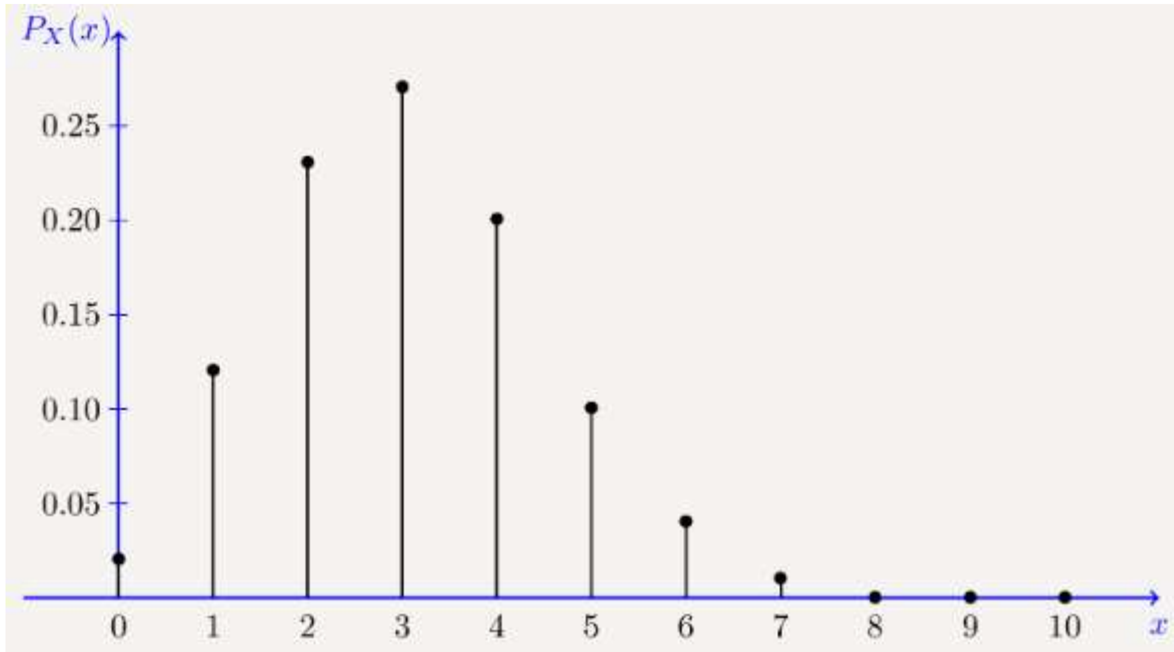
Лако је проверити да ове вероватноће, P_k, чине расподелу, Бернулијеву расподелу. Наиме, ако се десило k исхода онда се није десило неких j исхода ($j \neq k$), а додатно збир свих је један:

$$\sum_{k=0}^{n} P_k = \sum_{k=1}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$$

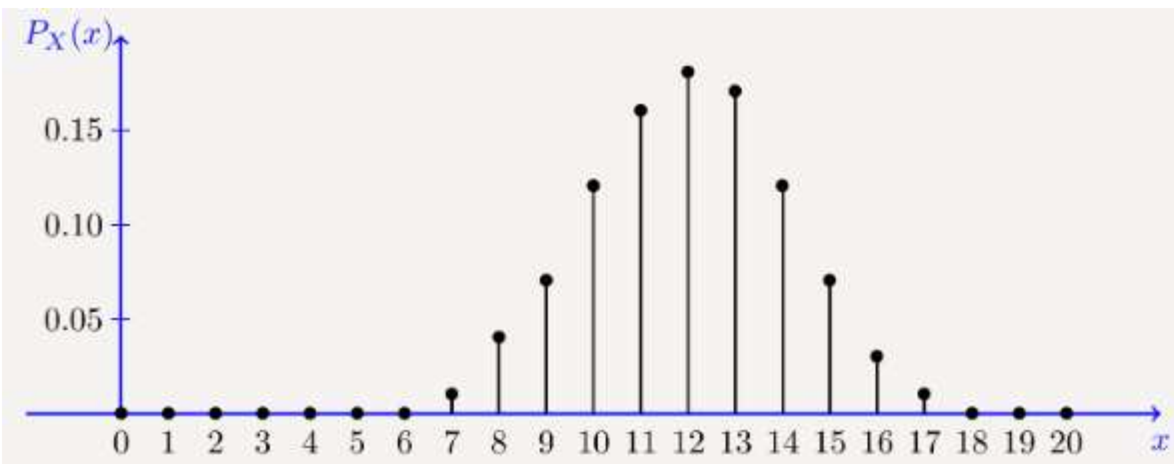
Доказ да је очекивање ове расподеле $\sum_{k=0}^{n} kP_k = np$ погледајте у теореми 2.3.1. књиге „Физичка информација“, да је варијанса $\sigma^2 = npq$ у теореми 2.3.3, а физичка информација $L_n = nL_1$, где је $L_1 = -p \log p - q \log q$ информација само једног таквог опита. ❑

Дефиниција 1. Случајна варијабла X има биномну расподелу $B(n, p)$, са параметрима $0 < p < 1$ и $n = 1, 2, 3, \dots$, ако је њена функција вероватноће дата са $P_X(x) = \binom{n}{x} p^x q^{n-x}$, где је $x = 0, 1, \dots, n$. ▨

У истој књизи наћи ћете и апроксимације управо дефинисане Бернулијеве, тј. биномне расподеле у ужем смислу (пример 4), са Пуасоновом и Гаусовом, за које даље следи неколико примера ради лакшег разумевања њихове повезаности. На слици је $B(10, 0.3)$, датог броја понављања $n = 10$ и вероватноће повољног исхода $p = 0.3$, дакле неизгледном.



На графу се види да је $x = 3$ најчешћи број жељених исхода у низу од $n = 10$ покушаја, при шанси $p = 0.3$ да се такав деси, односно шанси $q = 0.7$ да се такав не деси. То је некако и очекивано тзв. „очекивање“, $E(X) = np = 3$, које је горе поменуто за Бернулијеву расподелу. Просек квадратног расипања око средње вредности, $\sigma = \sqrt{npq} = \sqrt{10 \cdot 0.3 \cdot 0.7} = \sqrt{2.1} \approx 1.45$, такође је видљив.

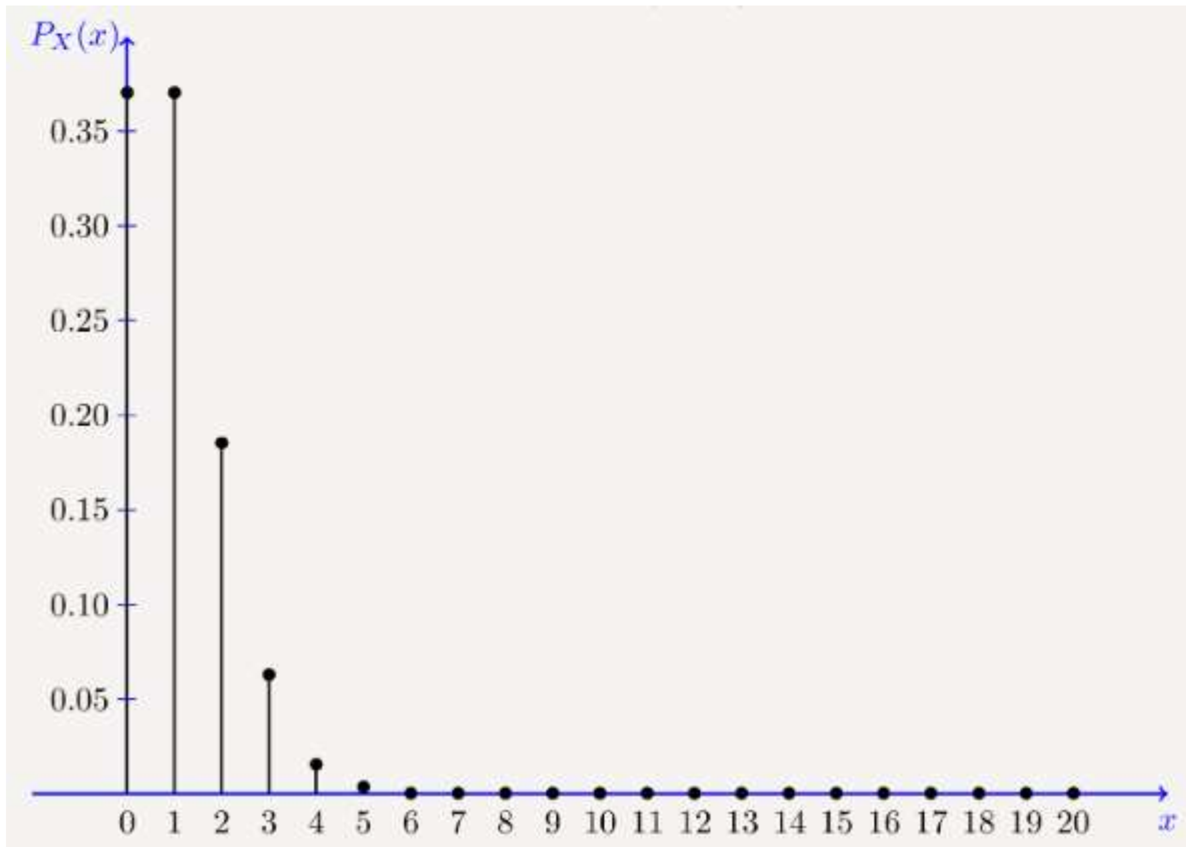


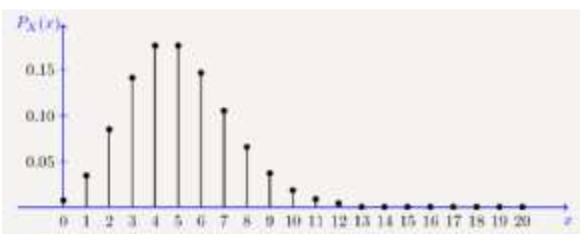
На другој слици приказана је расподела $B(20, 0.6)$, где се у низу од $n = 20$ покушаја, вероватноће $p = 0.6$ најчешће налазило $x = 12$ жељених исхода, такође очекивано, $E(X) = np = 20 \cdot 0.6 = 12$. Дисперзија око ове средње вредности је $\sigma = \sqrt{20 \cdot 0.6 \cdot 0.4} \approx 1.55$.

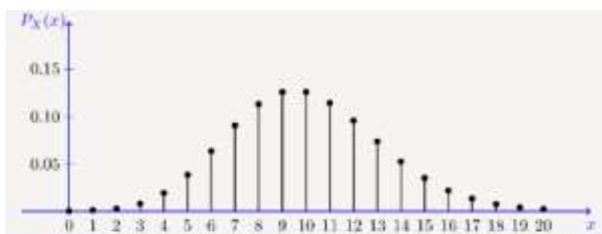
Средња, иначе максимална вредност графа помера се у десно са повећањем вероватноће p. Када је тај максимум (np) већи, дисперзија (npq) је мања, јер је укупни збир свих ових „стубића“ један. Како је дисперзија парабола, $np(1 - p)$, са гранама окренутим доле и нулама у тачкама 0 и n, то је њен максимум $n/2$, који се постиже параметром $p = 0.5$. То објашњава зашто и када Бернулијеву расподелу апроксимирамо Пуасоновом и Гаусовом.

Дефиниција 2. Случајна варијабла X има Пуасонову расподелу $P(\lambda)$, са параметром $\lambda = np$, ако је њена функција вероватноће дата са $P_X(x) = e^{-\lambda}\lambda^x/x!$, где је $x = 0, 1, \dots, n$. ▨

На првој од следећих слика је Пуасонова расподела са параметром $\lambda = 1$, на другој са параметром $\lambda = 5$, а на трећој са параметром $\lambda = 10$, редом лево и десно. Овај параметар λ, такође је средња вредност биномне расподеле, $B(n, p)$, тј. њено очекивање ($\lambda = np$).







Примере сам узео са сајта Special Distributions, где се могу наћи детаљнији описи и објашњења, па их (сличне) не морам овде пописивати. То су школски примери које можете налазити и другде. А битно и овде довољно је да приметите приближавање облику „Гаусовог звона“, ове, Пуасонове расподеле, у случају када се њен максимум, параметар λ, помера ка средини (домена апсцисе). Тај параметар је и средња вредност и варијанса Пуасонове расподеле, $\lambda = E(X) = \mathrm{Var}(X)$. Што је врх више десно – нижи је, а граф ове расподеле пљоснатији је.

35. Апроксимације

Под насловом „2.3.1 Апроксимације“ у књизи „Физичка информација“ доказан је лимес

$$\lim_{n\to\infty}\binom{n}{k}p^k q^{n-k} = \frac{e^{-\lambda}\lambda^k}{k!}$$

где је $\lambda = np$, када је $\lambda < 10$. То су ситуације малих вероватноћа, рецимо у протоку догађаја током времена какви су звоњава телефона, наилазак аутомобила, космичких зрака. Ако су они хомогени (расподела догађаја зависи само од трајања), независни и сепарабилни (не дешавају се два у исто време), онда имамо ситуацију за Пуасонову расподелу.

Пример 1. (Број посета) Нека је веб локација примала у просеку 10 посетилаца на сат. За $\lambda = 10.0$ и $x = 3$ у кумулативним таблицама Пуасонове расподеле налазимо $P(X \le x) = 0.0103$, што значи да ће током једног сата тај сајт имати не више од 3 посетиоца са вероватноћом 0.0103. Исто каже и кумулативни калкулатор Пуасонове расподеле. ❏

У књизи је поменути лимес доказан, а за следећу је речено је да се изводи слично. Комплетирања ради за ту следећу, Гаусову апроксимацију, ево Муавр-Лапласових доказа.

Теорема 1. У Бернулијевој расподели $B(n,p)$, када $n \to \infty$ тада за вероватноћу приближно важи

$$P\{X = k\} \sim \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-x^2/2}, \text{ где је } x = \frac{k-np}{\sqrt{npq}}$$

униформно по x у сваком коначном интервалу.

Доказ: Из $k = np + x\sqrt{npq} \to \infty$ и $j = n - k = nq - x\sqrt{npq} \to \infty$, кад $n \to \infty$, јер је x у коначном интервалу. Стирлингова формула $n! \sim \sqrt{2\pi n}n^n e^{-n}$ даје:

$$P\{X = k\} \sim \frac{\sqrt{2\pi n}n^n e^{-n}}{\sqrt{2\pi k}k^k e^{-k}\sqrt{2\pi j}j^j e^{-j}} p^k q^j = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{kj}}\left(\frac{np}{k}\right)^k \left(\frac{nq}{j}\right)^j$$

Поред тога,

$$\frac{kj}{n} = n\left(p + x\sqrt{\frac{pq}{n}}\right)\left(q - x\sqrt{\frac{pq}{n}}\right) \sim npq$$

па имамо

$$P\{X = k\} \sim \frac{1}{\sqrt{2\pi}\sqrt{npq}}\left(\frac{np}{k}\right)^k \left(\frac{nq}{j}\right)^j$$

Помоћу $\log(1+t) = t - \frac{t^2}{2} + O(t^3)$ налазимо:

$$\log\left(\frac{np}{k}\right)^k \left(\frac{nq}{j}\right)^j = -\left(np + x\sqrt{npq}\right)\log\left(1 + x\sqrt{\frac{q}{np}}\right) - \left(np - x\sqrt{npq}\right)\log\left(1 - x\sqrt{\frac{p}{nq}}\right) =$$

$$= -\frac{x^2}{2} + O\left(\frac{1}{\sqrt{n}}\right)$$

што је и требало доказати. ■

Интегралећи ову апроксимацију (сабирајући инфинитезималне кораке) налазимо следећу.

Теорема 2. У расподели $B(n, p)$, када $n \to \infty$, за интегралну вероватноћу важи

$$P\left\{a \le \frac{X - np}{\sqrt{npq}} \le b\right\} \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx$$

Доказ: Редом за $n = 1, 2, 3, \dots$ и $k = 0, 1, \dots, n$ пишемо:

$$x_{n,k} = \frac{k - np}{\sqrt{npq}}$$

$$x_{n,k+1} - x_{n,k} = \frac{1}{\sqrt{npq}}$$

$$P\left\{a \le \frac{X - np}{\sqrt{npq}} \le b\right\} = \sum_{k: a \le x_{n,k} \le b} P\{X = k\} \sim$$

$$\sim \frac{1}{\sqrt{2\pi}} \sum_{k: a \le x_{n,k} \le b} \frac{1}{\sqrt{npq}} e^{-\frac{x_{n,j}^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx$$

чиме је теорема доказана. ■

Сам резултат, апроксимација биномне расподеле $\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx$ је позитиван број, који достиже 1 са границама $a = -\infty$ и $b = +\infty$, па се може посматрати и као вероватноћа расподеле

$$P\{a \le Z \le b\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx = \Phi(b) - \Phi(a)$$

а може се доказати (Физичка информација, Теорема 2.4.1.) да је $\rho(x) = e^{-\frac{x^2}{2}}$ густина расподеле. А расподела је нормална, $N(\mu, \sigma^2)$, или Гаусова, са очекивањем $\mu = np$ и варијансом $\sigma^2 = npq$.

Таблице стандардне нормалне расподеле дају вредности функције

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{x^2}{2}}\, dx$$

које се могу добити и разумети и помоћу калкулатора нормалне расподеле. Примена је лака.

Пример 2. Код 24 400 бацања фер новчића писмо је пало 12 428 пута. Тражимо вероватноћу тог и већег одступања броја падања писма од очекиваног $np = 24400 \cdot 0.5 = 12\,200$, где је дисперзија

$\sigma = \sqrt{npq} = \sqrt{24400 \cdot 0.5 \cdot 0.5} = \sqrt{6\,100} \approx 78.102$. Нека је X број који показује број падања писма у $n = 14400$ бацања новчића, а ми тражимо вероватноћу догађаја:

$$P\{12\,428 \le X < \infty\} = P\{228 \le X - 12\,200 \le \infty\} =$$

$$= P\left\{\frac{228}{78.102} \le \frac{X - 12\,200}{78.102} \le \infty\right\} = P\{2.9 \le Z \le \infty\} \approx$$

$$\approx \frac{1}{\sqrt{2\pi}} \int_{2.9}^{\infty} e^{-\frac{x^2}{2}}\, dx = \Phi(\infty) - \Phi(2.9) =$$

$$= 0.50\,000 - 0.49\,813 = 0.00\,187$$

а то је прилично мала шанса. ❑

Приметимо да се исти шаблон може применити за добијање овде раније доказаног [26. Пример 1] Бернулијевог закона великих бројева, сада на другачији начин.

Пример 3. (Бернулијев закон великих бројева) Када $n \to \infty$, тада релативна учесталост x/n постаје све ближа вероватноћи p. То проверавамо. Овде је x број реализованих жељених исхода, а n број свих опита, док је p вероватноћа датог исхода. За свако $\varepsilon > 0$ биће:

$$P\left\{\left|\frac{x}{n} - p\right| \le \varepsilon\right\} = P\left\{-\varepsilon \le \frac{x}{n} - p \le \varepsilon\right\} =$$

$$P\{-n\varepsilon \le x - np \le n\varepsilon\} = P\left\{\frac{-\varepsilon \cdot n}{\sqrt{npq}} \le \frac{x - np}{\sqrt{npq}} \le \frac{\varepsilon \cdot n}{\sqrt{npq}}\right\}$$

$$= P\left\{-\varepsilon \cdot \sqrt{\frac{n}{pq}} \le Z \le \varepsilon \cdot \sqrt{\frac{n}{pq}}\right\} \approx \Phi\left(\varepsilon \cdot \sqrt{\frac{n}{pq}}\right) - \Phi\left(-\varepsilon \cdot \sqrt{\frac{n}{pq}}\right)$$

$$= 2\Phi\left(\varepsilon \cdot \sqrt{\frac{n}{pq}}\right) \to 1$$

када $n \to \infty$, јер су ε, p, q фиксирани бројеви, а $\Phi(\infty) = \frac{1}{2}$. Дакле, са повећавањем броја опита n, вероватноћа да би разлика релативне учесталости $\frac{x}{n}$ и вероватноће p могла премашити унапред дати, произвољно мали број $\varepsilon > 0$, тежи нули. ❑

Пример 3 појашњава зашто 2. пример даје вероватноћу од само 0.00 187. Та „неочекивано“ мала вероватноћа за скроман број од 228 или више пребачаја средње вредности, након читавих 24 400 покушаја, може нам се учинити значајно премаленом. То је зато што живимо „хоризонтално“, или окомито на „вертикалу“ деловања закона великих бројева, када је разумемо као линију повећања из микро у макро свет физике. Ми нисмо грађени да лако владамо правцима промена настајалих преласцима из малобројног у многобројно.

Да није тако не би хипотеза да физички не постоји нешто малено попут атома, оно што не можемо видети или осетити (Ернст Мах) нашим чулима, уопште била прихваћена и не би била тако жилаво

брањена у време самог почетка 20. века. Не бисмо физиком називали само оно што је непосредно обзервабилно да прозор чула која имамо не показује довољно света. Али ту је и математика да да предност својој специфичној логици наспрам опште прихваћених веровања, па ипак разматрамо и „интуитивно неприхватљиве“ ставове.

У случају великих бројева, фреквенције достигнуте кроз дуги низа покушаја, формуле вероватноће као да говоре о некој врсти „памћења“. Природа узима у обзир и прошлост у оцени садашњости, а подсећам, то се види не само у теорији вероватноће. Пратећи положаје звезда на небу, древни су народи знали када почиње нова година. Начин хода, дневне рутине као и потпис, карактеришу нас појединачно. Често смо у стању процењивати шта ће бити на основу онога шта је било, а модел 3. примера на неке начине говори нам да дугим ходом времена (садашњости) те процене постају све извесније.

Из огољеног закона великих бројева, самог по себи, онаквог како га види „чиста математика“, не следи и развој физичких система у вероватније. Природне су појаве сложеније. Закони о великим бројевима случаја кажу да понављањем „мерена учесталост постаје све ближа вероватноћи“. Они су необични колико и раст запремине са кубом висине наспрам пораста површине са квадратом и успореност, или трапавост, крупних животиња наспрам релативној брзини (у односу на њихов свет величина) и живахности малих. Неке разлике између светова малог и великог, а налик 2. примеру, можа даље препознајете и сами. Оне су нам необичне колико и неистражене.

Вероватније јесте мање информативно, па ићи ка извеснијем значи ићи ка мање информативном. У укупној супстанци космос тако постаје мање информативан, мање неизвестан. То је смисао тока ентропије из мање у већу, тзв. термодинамичке смрти васионе. Међутим, топећи се из супстанце информација прелази у простор и васиона буквално постаје све већа, све ређа супстанцом.

То није новост за „теорију информације“, иако ће бити за науку. Али у тој „старој“ причи новост је да у закону великих бројева постоји шаблон за сличне налазе (надам се да их има више). Типично математички апстрактни модел са каквим радимо када сабирамо „два плус два је четири“ једнако рачунајући јабуке као и километре, иако знамо да су ти објекти физикално веома различити, треба наћи и развити[8]. Чак и појмовно, па и апстрактно, извесно је да одређени различити објекти могу имати исте математичке форме.

Велики математичари знали су описати математику као „вештину давања истих имена различитим појавама“. Са акцентом на вештини, ту карактеристику „препознавање истог у различитом“ треба разумети као тежину умећа или посла откривања математике.

[8] Савет мени за касније.

36. Расипање

Најопштији облик густине нормалне $N(\mu, \sigma^2)$, или Гаусове расподеле ($e = 2.71828 \dots$) је

$$\rho(k, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-\mu)^2}{2\sigma^2}}$$

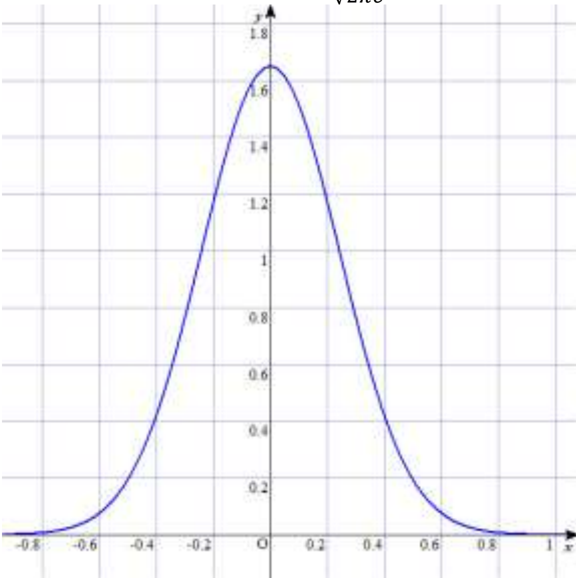
где су μ (математичко) очекивање, а σ^2 варијанса [35. Теорема 1]. Промењљива k обично је x. У Информацији Перцепције, у наслову „3.2 Информација“ израчунавао сам

$$\langle I \rangle = -\int_{-\infty}^{+\infty} \rho(x, \mu, \sigma) \ln \rho(x, \mu, \sigma) \; dx = \ln\left(\sigma\sqrt{2\pi e}\right)$$

што је средња вредност Хартлијеве информације ове густине, односно Шенонова информација. То је резултат који је онда било тешко пронаћи у литератури, за разлику од данашњих где се он јавља свукуда. Важан нам је да приметимо да се информација нормалне расподеле, након интегрирања, сабирања свих инфинитезималних корака, опет приказује у облику Хартлијеве.

Средња вредност континуалне информације, $\langle I \rangle$, која апроксимира бинарну у случају великог низа понављања, није баш тачно једнака Хартлијевој. Прво, густина нормалне расподеле није n једнако вероватних исхода, да би Хартлијева информација износила $\log n$, са избором базе логаритма који дефинише јединицу мере информације. Али знамо да се такав број увек може мењати фиктивним истим, замењен „константним средњим“ вредностима.

Друго, са дисперзијом $\sigma < \frac{1}{\sqrt{2\pi e}} = 0.24197$ тај логаритам постаје негативан, а на слици лево граф



је граничног случаја $\sigma\sqrt{2\pi e} = 1$, где је $x = \sigma$, а средња вредност $\mu = 0$. За дисперзије мање од ове, нормална расподела личи на троугаону[9] и, наравно, са све вишљим шиљком.

Међутим, знамо да за n велико, рецимо $n > 50$ а мало p, уз $np < 10$, Бернулијева (бинарна) се расподела своди на Пуасонову. Иначе тачнија је Гаусова (нормална) расподела. Апроксимација Бернулијеве Гаусовом даје:

$$\sigma^2 = npq = np(1-p)$$

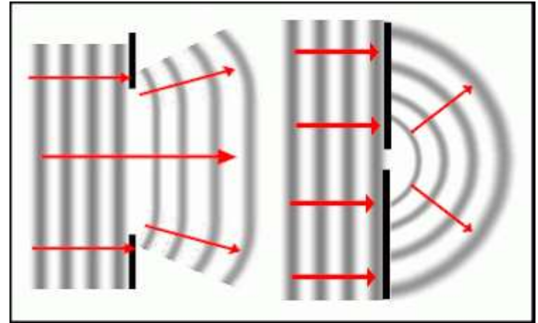
а мале вредности се постижу за $p \to 0$, или пак $p \to 1$. Према томе, те мале дисперзије онда и нису за ову, него за Пуасонову расподелу.

Треће ограничење је сама континуалност. Она се опрезно и са резервом треба користити у крајностима, јер сви наши опажаји коначни су. Скупа, сва три ограничења нас упозоравају да не треба журити са обрнутом применом, генерализацијом просте Хартлијеве информације [05.].

[9] Р. Вуковић: *Физичка Информација*, Економски Институт Бања Лука, 2019. наслов „Троугаона расподела“.

Међутим, расипање о којем ћемо овде говорити не тиче се само нормалне расподеле. Занимљив нам је и Комптонов ефекат (Scattering). Као што можемо видети у „теорији информације“ (не и ван ње), након судара фотона са електроном (Compton Effect), фотон један део своје енергије предаје електрону, тако да се енергија фотона смањује, а таласна дужина му се повећава.

Код обичне дифракције, на слици десно, таласи скрећу не мењајући таласне дужине, осим када се ради о проласку срединама кроз које се талас креће промењеном брзином. Тако ће у „гушћој“ средини, мање брзине, таласне дужине бити краће, а у већој брзини и таласне дужине бити веће (Снелов закон).



У Комптоновом ефекту скрећући фотон добија веће таласне дужине, али брзина светлости у вакууму не мења се, те се светлост ту не показује као талас већ као честица (корпускула). Одавно тој појави додајем „информатичко“ тумачење. Одбијајући се од електрона фотон прелази на путању увећане таласне дужине и смањене енергије. Положаји му тада постају више размазани, неодређенији.

Физика се ту зауставља, нема „више размазаних“, али информатичка теорија разматра ситуацију и даље, прво као налаз да се актуелне путање могу посматрати и као највероватније (са становишта честица-таласа које се крећу). Затим и као насилно скретање фотона (сударом) на мање вероватну трајекторију, што је налаз отпора већој неизвесности, односно присуства спонтаног развоја стања у мање информативна. Треће, утапање у неизвесност компензује додатну брзину фотона.

Фотон на неизвеснију путању не иде спонтано. Тај прелазак спречавају „силе неизвесности“, а тек друге силе (судар) то му скретање омогућују, надвладавајући отпор. Додатно примећујемо, фотон као да тоне у неизвесност попут „истине“, односно „реалности“ која слично тако бледи, скрива се, постајући „лаж“, односно која одлази у „псеудо реалност“ (The Truth). То потискивање потпомаже привлачност разређене информације, а онда неке исте математичке форме крију се чак и у таквим двема различитим појавама.

Скривање у некој лажној реалности заправо виђамо свугде око себе. Оно је у популарности лажи, фикције или романа већој од доказа математичких ставова и сувопарних истина, па до испадања из исте садашњости удаљених делова истог тела (светлости треба неко време да пређе пут). Али и прошлост је врста псеудо-реалности. Још екстремнија таква врста је део сопственог покретног тела који релативни посматрач не може непосредно виђати, па оно што може опажа као успоравање тока времена (количина догађаја дефинише време).

Неке од тих феномена остављам недовршеним, за сада, јер имамо више опција од чврстих места. Алтернативне идеје и теорије и даље су могуће, а верујем те ствари су још компликованије. Није ни Мендељеј наша систем елемената онако једноставним како су га древни народи замишљали (земља, ватра, ваздух, вода), нити су атоми испали тако просте грађе каквој су се надали (Леукип и Демокрит), па не журим.

Елем, фотон спречен контактом са електроном да одржи стару путању, рецимо „убеђивањем“, или комуникацијом, разменом информација, да не кажем само интеракцијом, фотон преузима путању већих таласних дужина. Доследно теорији информације, било би да електрон наставља

путовање на трајекторији мањих густина вероватноћа. Са становишта „сила неизвесности“ [27.], фотон тамо иде присилно, јер размазане позиције којима даље корача мање су изјашњене, неодређеније су. Посебно је питање, како фотон опстаје у том одбојном окружењу?

Свет се састоји од реалних и псеудо-реалних информација. Када бисмо прве назвали јасним, или апсолутним истинама, ове друге биле би потиснуте истине или лажи. Истине су јаче, аутентичне реалности, лажи су привлачније (The Truth). Лажи делом смање одбојност неизвесности којој иду фотони након расипања око електрона. Дакле, фотон скретањем на веће таласне дужине и мање енергије након одбијања од електрона делом постаје „потиснута реалност“, нејачи у смислу блеђи у односу на бившег и утопљен у већу неодређеност положаја. Видљива енергија и импулс фотона предата електрону, тачно су једнаки енергији и импулсу електрона предатих фотону, а псеудо део фотона личи на кодирану, потиснуту, или прикривену истину у лажи.

Теорија информације предвиђа различите врсте сличних „потискивања“. Супротан пример нејакој али привлачној лажи је непривлачан систем у релативном кретању у којем се енергије виђају веће (за кинетичку енергију). Део покретног система видљив само сопственом посматрачу је у псеудо реалности за релативног, па овај други који опажа кретање може само претпостављати постојање додатне информације, тако опажати неизвесност и не прелазити у то стање кретања. Ово је сада објашњење са становишта у претходним насловима дискутованих „сила неизвесности“.

Подсећам, тај вишак информације релативно опажан значи и мањак релативне ентропије и опет одбојност напуштања стања мировања (у свом систему) релативног посматрача. Закон инерције према томе следи из вишка псеудо-реалности коју тело вуче за собом, а према овом објашњењу, из отпора који псеудо-вишак ствара као „опажена“ (секундарна, претпостављена) неизвесност.

Тело у јачем гравитационом пољу виђено из слабијег такође има спорији ток времена. По истој логици (да количина догађаја дефинише време), оно је делом у псеудо реалности. Гравитација закривљује 4-Д простор-време у неком ширем бар 6-Д универзуму (са две димензије времена вишка). Оно зато има увећану масу и енергију – са становишта релативног посматрача, јер му је зато лако претпоставити присуство дела вишка тела у псеудо реалности. Неизвесност заправо и није другачије „видљива“ сем посредно.

Одбијање тела од тела није исто што и Комптоново расипање фотона око електрона. То пре свега због константне брзине светлости, а даље јер фотони немају сопствено време. Они не могу бити „делом у псеудо реалности“, због чега и као одбијени таласи губе енергију као целина. Тела имају вишак времена и неизвесности са којим могу додатно манипулисати.

37. Репрезентативан узорак

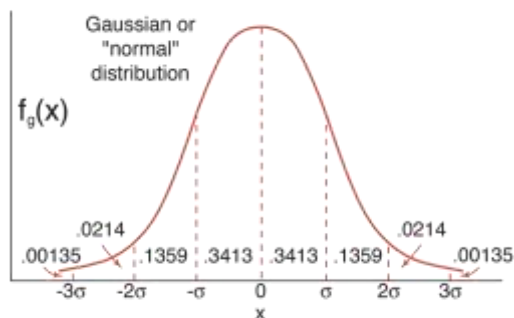
Необична је и интересантна сама идеја о „притиску неизвесности“ [27.]. Она ненадано доследно додаје неке формализме физичких „сила“ области вероватноће. Потврђује објективну случајност и концепт силе у теорији информације, макар на крњем нивоу каквом га данас видимо. Занимљиво је проверити пружа ли се она тако широко кроз насумичност, као и у динамици.

У време када сам писао „Информацију Перцепције“, (Економског института Бањалука, 2016) само сам поменуо тада познат „Проблем секретарице“, али тек у Прилозима II (почетком 2021), наслов „1. Репрезентативан узорак“, садржавао је нешто од дубљих мотива и јаснији циљ те теме. Стара тема, нова допуна.

(Проблем секретарице: На располагању је једно радно место за секретара/секретарицу. Комисија зна број подносилаца захтева и интервјуише их случајним редоследом, једног по једног, не знајући ко је следећи. Рангирање заинтересованих довољно је детаљно да не буде значајног дуплирања скора најбољих, а одлука о избору заснива се само на претходним резултатима. Након одбијања текућег кандидата касније га није могуће позивати, а након прихватања даља претрага се обуставља. Како изабрати најбољег кандидата?)

Грубо речено, ради се о трећини случајно расутих исхода који „некако“ постају репрезентативан узорак. То решава „секретарицу“, али уноси и појашњење облика ,,Гаусовог звона''. Трећина графа садржи углавном све битно за расподелу, а необично је што управо чиста случајност гура довољну главнину около средње вредности. Додатно, граф иде тако „незгодном“ линијом да ће покушаји симулације неизвесности, рецимо напамет, из главе наводећи „случајне“ бројеве, па и бољим наштимавањем резултата (лутрије), њену линију значајно промашивати.

Тестови случајности (теорије вероватноће) заснивају се на тој незгоди кривине „Гаусовог звона“. Скоро је немогуће имитирати је, тако да и тестови за откривање евентуалних превара у играма на срећу више вреде. Гаусова крива, на слици лево, показује још једну општу особину. Област унутар дисперзије $\pm\sigma$, колика год она била, покриваће око трећине (± 0.34) иоле важног дела густина. Стојећи на становишту „сила неизвесности“ рекли бисмо да око 2/3 неизвесности уравнотежује 1/3 извесности. То би могао бити централни део оваквих прича.



Међутим, нормална расподела са своје стране има апроксимацију. Развој експоненцијалне функције у ред и примена на густину вероватноће, даће нам првих неколико њених сабирака значајних за укупни збир:

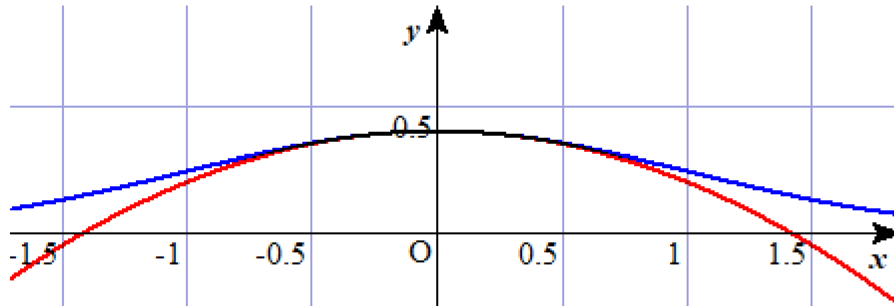
$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots$$

$$y = \rho(x, \mu = 0, \sigma = 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2^k k!} = \frac{1}{\sqrt{2\pi}} \left(1 - \frac{x^2}{2} + \frac{x^4}{8} - \frac{x^6}{48} + \cdots\right)$$

То је (приближно) парабола

$$y \approx \frac{1}{\sqrt{2\pi}}\left(1 - \frac{x^2}{2}\right)$$

У околини нуле ($x \approx 0$) довољно је узети прва два сабирка развоја Гаусове густине, јер су четврти и виши степени x-а тада занемарљиви. Граф постаје парабола, доња црвена линија на слици. Линија изнад (плава) је граф тачних вредности густине. Око нулте, средње вредности, поклапање ова два графа је значајно, што значи да у том домену оба опадају са квадратом апсцисе.



Приметимо да силе које опадају са квадратом удаљености већ постоје у физици (гравитациона и Кулонова), да су конике путање набоја које такве покрећу (као расподеле у наслову [28.]), а да им се дејства простиру брзином светлости (што се подразумева и за случајне исходе).

Са овим се не исцрпљују аналогије физике сила и неизвесности. У књизи „Простор-Време“, унутар наслова „1.2.8 Вертикалан пад“ наћи ћете извођење $m = m_0 e^{GM/rc^2}$, где је m релативна маса тела које пада у гравитационо поље масе M, када је оно на удаљености r од центра, m_0 сопствена маса тела, G је гравитациона константа, а c брзина светлости. Даље се може видети да је овај експонент потенцијална енергија, која се у слободном паду претвара у кинетичку, а да једначина подељена квадратом брзине светлости, због $E = mc^2$, постаје однос енергија тела. Пређимо на фреквенције и налазимо локално трајање једне фазе таласа енергије $\tau = \tau_0\, e^{-GM/rc^2}$ у односу на, τ_0, сопствено трајање.

Сопствено трајање је врста константе, јер представља универзални период, па је последњи израз формално једнак густини Гаусове расподеле. Наиме, успоравање тока времена функција је броја исхода, а кинетичка енергија се мења са квадратом брзине v. Ова у експоненту као промењљива ($x = v$) даје управо тај незгодни експонент Гаусове формуле. Са друге стране, оваква каква је, та формула, о успоравању релативног времена у гравитационом пољу, представља и везу између вероватноће и (Хартлијеве) информације.

Толико за сада о „чудном“ питању, да ли се у гурању случајности ка средњој вредности можда не сакрива и још понеки непознати формализам физике сила. Ти формализми су нам познати, што се тиче теорије информације барем. Они постоје у спонтаном развоју мање информативних стања, у току ка мањој комуникацији, или мањем дејству. Изражавају управо „принцип најмањег дејства“, одавно познат, из којег теоријска физика изводи све, али баш све нама данас важеће трајекторије, иначе настајале под дејствима сила.

38. Виртуелне сфере

У наставку „одбојну силу“ неизвесности подразумевамо, а како о њој мало знамо претпостављамо неке њене физикалне карактеристике. Пре свега то је неки закон одржања (Conservation) који тек требамо откривати. Акценат је на апстракцији коју бисмо (можда) могли наћи у конкретном.

За почетак приметимо да неизвесност врши притисак на извесност, те да се ова друга држи попут мехура од сапунице тежећи концентрацији, току ка мање информативном стању. Међутим, каква год била вероватноћа да се догађај деси ($0 < p < 1$), увек постоји вероватноћа да се исти не деси ($q = 1 - p$), што ће рећи да о ма каквој количини неизвесности причали, у њој постоји и количина извесности. Представисмо речено сфером (мехуром) са гушћом унутрашњом извесношћу.

Разматрамо раније постављено питање о равнотежи извесности и неизвесности, прве унутар дате области (сфере), а друге изван. Говоримо, наравно, о апстрактним силама, скупљања и ширења, а које дефинишу границу између две области. Да нема прве, тежње ка мањој информацији и према томе концентрације ка унутрашњости сфере, надвладала би друга, вањско гурање од себе, тако да граф нормалне расподеле вероватноћа успева бити звонаст, не претвара се у уски вертикалан стуб на средњој вредности.

Промене ових утицаја, концентрације против надирања, на слојевима количина уравнотежују се зависно од садржаја. Када концентрација извесности са даљином r од свог центра опада, тада сав систем може ући у равнотежу тако да је на свакој од тих концентричних сфера исти притисак обеју страна. То је може рачунски показати у различитим метричким системима, а није тешко разумети ни интуитивно у еуклидском.

Наиме, када је једном успостављен систем опадања неизвесности са удаљеношћу од центра, онда сфера све већег полупречника може затварати све више извесности, а на себи имати мању и мању њену густину по јединици површине. Разлика између густине неизвесности дела сфере и вањског фронта може бити константна. Једно од решења тог задатка управо су сфере обичне геометрије са површинама које расту са квадратом полупречника ($4r^2\pi$), а тада густине дела површине опадају са квадратом удаљености. Резултат је заправо једнак ономе који смо расправљали у претходном наслову [37.]. Имам и незнатно другачији приступ истом, али прво видимо шта је са дејством.

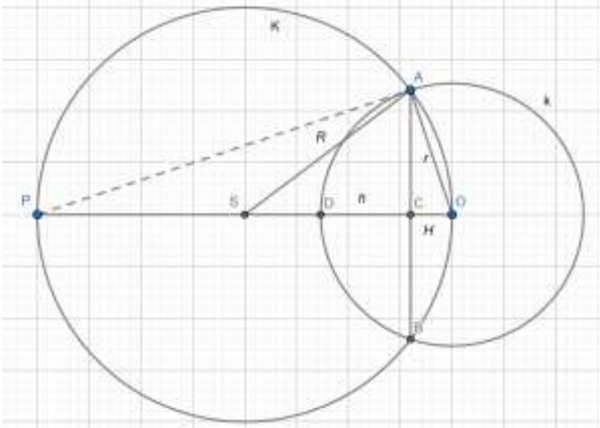
Брзину светлости ($c \approx 300\ 000$ km/s) требамо узети у обзир, јер се ове статичне сфере шире том брзином у односу на (произвољног) посматрача. Наиме, оне су безвременске, посматрано у само једном тренутку (било које поједине) садашњости, али садашњост клизи ка будућности брзином светлости. То је један од ставова теорије информације, објаснићу.

Простор-време је најмање 6-димензионално. Од шест, три димензије су просторне, а следеће три су временске. Све једначине физике, према овој теорији, требале би бити инваријантне на изборе четири од таквих шест димензија, од којих су три просторне координате (попут x, y, z) а четврта је време, оно множено имагинарном јединицом ($i^2 = -1$) и брзином светлости даје дужину. Опште једначине релативности већ су такве, а и релативистичка Клајн-Гордонова квантне механике. Сви објекти који немају сопствено време, попут фотона, добијају га од посматрача. Према томе, сви се такви „крећу“ брзином светлости.

Подсећам, овде је важно нагласити да квантна спрегнутост долази од истовремености, такође у званичној физици још увек непознат детаљ. Ове покретне сфере тако постају виртуелне, тренутне, или истовремене, за било којег посматрача, те су објективне и квантно спрегнуте.

Остатак света, ван површи сфере, у односу на произвољног посматрача није квантно спрегнут и подлеже ширењу дејства не брже од светлости. У следећем примеру анализирана је таква једна велика сфера са центром S на чијој је површини центар O много мање сфере. Доследна горњој, сада је претпоставка да унутрашњост веће сфере у мањој има једнак утицај као и спољашњост, иако је запремина унутрашњег дела мања од спољашњег.

Пример 1. Дата је лопта L са центром S и полупречником R и лопта l са центром O на сфери прве и полупречника r. Слика лево је пресек лопти и равни која садржи центре ових сфера. Тако оне постају кружнице K и k, редом тих лопти, већа и мања. Пресеци кружница су две тачке $A, B \in K \cap k$, пресек већег полупречника SO и мање кружнице је тачка D. Пресек тетиве AB и полупречника OD је тачка C.



Тако је $R = \overline{SA} = \overline{SO} = \overline{SP}$, где је $\overline{PO} = 2R$ пречник већег круга, затим $r = \overline{OA} = \overline{OD}$, а $H = \overline{OC}$ и $h = \overline{CD}$. Троуглови POA и AOC су слични, па је $H = r^2/2R$, а иначе $h = r - H$.

Тражимо запремину дела мање лопте која припада већој. Она је збир два лоптина одсечка:

$$V_1 = V_k + V_K = \frac{h^2\pi}{3}(3r - h) + \frac{H^2\pi}{3}(3R - H)$$

Сређивањем и уврштавањем H и h лако налазимо:

$$\begin{cases} V_1 = \dfrac{r^3\pi}{12R}(8R - 3r) \\ V_2 = \dfrac{r^3\pi}{12R}(8R + 3r) \end{cases}$$

где је $V_2 = V_k - V_1 = \frac{4r^3\pi}{3} - V_1$. Однос запремина дела мање лопте која припада већој је:

$$f = \frac{V_1}{V_2} = \frac{8R - 3r}{8R + 3r} = \frac{1 - \frac{3r}{8R}}{1 + \frac{3r}{8R}} \approx \left(1 - \frac{3r}{8R}\right)^2 = 1 - \frac{3r}{4R} + \left(\frac{3r}{8R}\right)^2 \approx 1 - \frac{3r}{4R}$$

где је приближност утолико тачнија што је друга лопта мања од прве. ❑

Пресек веће и мање лопте има запремину V_1, а остатак мање лопте запремину V_2. Интуитивно, као и из резултата у примеру, јасно је да су ове запремине неједнаке, $V_1 < V_2$. Разлика тих запремина смањује се и исчезава када $R \to \infty$. Повећавањем веће сфере, мања се не мења ($r = \text{const}$). Мања представља опсег непосредног утицаја на околину тачке (O) на већој сфери, дејством које се шири константном брзином светлости.

Претпостављамо да је потенцијал, унутрашњег (V_1) и вањског (V_2) дела мање лопте у већој, једнак, па закључујемо да он опада са удаљеношћу. Колико пута је мања запремина прве толико пута је већи њен потенцијал, да би две силе биле у равнотежи. Промена потенцијала је на начин као код физичких сила, пре свега мислим на гравитацију (Простор-Време, 1.2.10 Шварцшилдово решење) и електро (Coulomb Force).

Овакве виртуелне сфере могуће је додати и Фејнмановим дијаграмима (Многострукости, Квантна спрегнутост). На пример, електрони их емитују као „виртуелне фотоне“. Њима се таласне дужине не мењају, а амплитуде су све мање. Прве би значиле пренос енергије, импулса и спина са једног електрона на други, а друге вероватноћу таквог преноса у ширењу сфере преко другог електрона.

Њихово ширење, са треће стране гледано, иде са изјашњених неизвесности према неизјашњеним, из унутрашњости лопти ка вани. Густина изјашњених (реализованих исхода) већа је унутра, што је налик кретању садашњости из прошлости ка будућности. Ово важи за произвољног посматрача, те је објективно. Замислимо ли посматрача којем време иде унатрашке у односу на наше, њему ће се ове сфере опет ширити. Сфере заправо и не иду, оне стоје, а „покреће“ их ток времена посматача.

Да би уопште постојале информације неопходно је време и објекти којима време не стоји. Наме, други пута саопштена „вест“ више није вест, па су и субјекти који нешто опажају увек у неком току времена. Они су ти због којих се садашњост креће ка будућности, посебно због којих се светлост креће. Према томе, аналогија развоја садашњости са кретањем виртуелних сфера има још дубље значење, значајније за физику него што бисмо помислили на први поглед.

Ми не можемо остајати у истој садашњости, неком 3-дим простору, просто зато што је развојност неопходна. Информација је ткиво простора, времена и материје, а неизвесност је њена суштина и зато ток догађаја иде ка њиховом изјашњавању. Упоређено са начелом шкртарења информације, ток догађаја нам говори о „наивности“ физичких система, о њиховом „несхватању“ да би тамо ка будућности можда могло бити више тога за избегавање него у прошлости. Или, пак, о „схватању“ да већ самим избором смера кретања оно где се иде постаје сиромашније информацијама.

Парафразирам и пренаглашавам да бих некако пренео поенту. Камо год да се догађаји развијају, увек се иде из прошлости у будућност, из гушћих ка ређим изјашњеним информацијама. То даје могућност развоју догађа ван нашег тока, па и „окомито“ (у линеарној алгебри ово је и буквално, а њена репрезентација је квантна механика) на њега, што је и неопходност уколико су случајности и опције објективне. Коначно, ово тражи и 6-дим простор-време где би избор четири димензије био произвољан, од којих би једна била време (множено са ic).

39. Примарна сила

У овом тренутку изгледа невероватно, али управо је тако, да се расподеле вероватноћа као и силе физике подвргавају истим законитостима. Ове прве, апстрактне математичке, невезане су за било које одређене физичке појаве, а друге, попут гравитационе, електро и свакако уопште замисливих константних централних сила. Иначе то успут излажем у овој скрипти „Информатичке Теорије“, да остане, да сутра можда неко примети како смо и ми данас досегнули те апстрактне висине.

Под насловом „38. Виртуелне сфере“, рецимо, наћи ћете да се Гаусова расподела вероватноћа понаша „као да“ постоји неки потенцијал „физичке силе“ који случајне исходе тера у граф познатог звонастог облика. Наводницима наглашавам да су ти потенцијали више од онога на што бисмо се лако навикли, да је њихова објективност изнад тога што бисмо на први поглед очекивали и да је она заједничко тло је за многе „праве силе“.

Под насловом „28. Центрифуга неизвесности“, у свом апстрактном простору варијабли, налази су да случајни исходи беже од неизвесности трајекторијама које су хиперболе. Иначе се по коникама (елипсе, параболе, хиперболе) крећу набоји вучени или гурани константним централним силама (привлачне по прве две, одбојне по трећој).

Под насловом „24. Примери регресије II“ налази се потврда да случајни исходи следе Кеплеров други закон (потег сунце-планета у једнаким временима пребрише једнаке површине), на начин који ће се показати да је више од аналогије.

Укратко, принципијелна тежња природе да чешће реализује вероватније исходе не само да на први поглед делује као нека (апстрактна) сила, већ је она то и на други и трећи поглед. Тачније речено, силе које сматрамо „физичким“ само су једна од врста тих апстрактних притисака, њихов израз, еквивалент, или последица. Доследно томе показаће се да једначине опште релативности можемо извести и из закона вероватноће, као што се могу извести из принципа најмањег дејства. Доказ овог другог погледајте у књизи „Минимализам Информације“, у наслову „2.5 Ајнштајнове опште једначине“.

Што је догађај вероватнији мања је вест, па „принцип максималне вероватноће“ (да се чешће догађају вероватнији исходи) тако постаје „принципом минималне информације“ (спонтано се системи развијају у мање информативне). Обадва даље сводимо на „принцип најмањег дејства“, иначе познат и признат у физици. Први од наведених (о вероватноћи) није нека нама чудна појава, штавише сви који су се бавили применом или развојем теорије вероватноће подразумевали су је, али не знам да је неко тако истицао и користио. Други (о информацији) је потпуно непознат[10].

Да ове „апстрактне силе“, које овде успут откривамо, нису споредна или локална појава може се демонстрирати и њиховим присуством у математичкој теорији игара. Погледајте прво мој кратки прилог „Win Lose“ (из фебруарског блога 2022) да бисте узели у обзир и разлике, тачније додатак моје теорије тој. Прочитајте затим и прилог „Loss The Winning Strategy“, или сличан, да овде скратим причу на детаље свакодневнице.

Ако теорија игара каже да стратегија „губитак-губитак“ (жртвом до победе) скоро увек побеђује „добитак-добитак“ (претерану оданост доброти), зашто ће се онда многи сложити да „Увек треба

[10] Добијам и добронамерне поруке од „познаваоца“ да тај принцип није тачан и да га избацим.

бити добар. Без обзира да ли ће вам се неко захвалити или не, памтити или не, да ли ће вам узвратити или не. Негде се све то пише.“ – Пита ме један и закључује – „Значи ниси у праву?“

Људи имају аверзију према губитку – одговарам. – Када мисле да им је добро, нерадо се подмећу и „не страдају без разлога“. Када зло надире пријатније је и даље веровати да ће то и само проћи, па ћемо рећи: „Спајај добро са добрим и лоше ствари ће се неким чудом решити без мене“. Затим је лако игнорисати ставове о потреби непријатног ангажовања, поготово зато што су ружне радње (управо за такве) смарајуће или прекомпликоване.

Ове примедбе цитирам са својих јавних објава (друштвених мрежа), одакле добијам и коментаре, критике или нова питања, понекад. Од таквих је и следеће: „Има ли негде још каквих прилога о нашој одбојности према ризику? Верујете ли у то? Имате ли своје лично објашњење?“

Да, заправо и те појаве истражујем. – Одговарам – Ево једног страног прилога „Why People Don’t Take Risks“ (превод: Живот је лак када се играмо малим. Лако је радити оно што сви око вас раде. Играти на сигурно, остати у кутији – чули сте клишее. Када не постанемо наши снови, заправо постајемо мањи...). Често нас многи убеђују како увек треба спајати добро са добрим, никада не отпором на непријатне иницијативе, а поготово не узвраћати злим на зло. Таквим заговарањима попуштања, ставова да увек треба ићи „као кладић низ воду“, не излазити из света свог комфора, не подилази (ако имаш снаге). Природа успеха је нешто друго.

Што се тиче трећег питања, имам тога додати, наравно. Неизвесност је одбојна сила, не само „као да је“ већ је она то и заиста. Отпор неизвесности, или рецимо не-вољење ризика, није само пука еволуцијом створена емоција, него је нешто што је еволуција покупила у емоције из физичког света. Али то је друга и дуга прича – кажем у том тексту.

Апстрактна природа „сила неизвесности“ појавиће се и у каснијим открићима, у наизглед далеким местима овим, јер је информација свеприсутна појава, она је ткиво света око нас и тако екстремна да је налазимо од најконкретнијих до најапстрактнијих појава и процеса које нашим перцепцијама можемо појмити, а има је и даље од тога.

Неизвесност постоји и осећамо је. То није исто као када (не)видимо неке електро магнетне таласе, који нам могу бити небитни. Информација је количина неизвесности, опција или физичког дејства и жива смо бића зависно од тога колико је имамо у односу на мртву твар од које се састојимо. Што смо виталнији даље смо од принципа најмањег дејства којег се доследно придржава неживи свет физике, али не изван удобне зоне наших емоција која је појас од минималне до максималне дате нам властите информације.

Принцип минимализма би да тај вишак информације одбацимо, али и околина је њоме попуњена и тежи истом, а закон одржања нас све заједно у томе омета. Зато нас плаши вишак неизвесности, јер он значи опције са којима се не умемо носити или тражи од нас додатну агилност. Изискује од нас неприродне радње, противне инертности, што на свој начин примећује и решава психологија (Dealing with Uncertainty). Са друге стране су потреба за слободом и страх од смрти.

40. Непрекидне перцепције

Знамо да информација перцепције $S = \vec{a} \cdot \vec{b}$ достиже максимум када су низови $\vec{a} = (a_1, \dots, a_n)$ и $\vec{b} = (b_1, \dots, b_n)$ исте монотоности, оба растућа или оба опадајућа. У крајњем другом случају, када је један растући а други опадајући, информација перцепције је минимална. Наводио сам то много пута [12. Екстреми], али углавном за дискретне случајеве (низове). Међутим, исте особине важе и за непрекидне перцепције, што се може видети из следеће теореме.

Теорема 1. Нека су f и g реалне, непрекидне функције, такве да је $\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} g(x)dx > 0$. Ако је интеграл $-\int_{-\infty}^{+\infty} f(x) \log g(x)\, dx$ коначан, тада $-\int_{-\infty}^{+\infty} f(x) \log f(x)\, dx$ конвергира и вреди неједнакост $-\int_{-\infty}^{+\infty} f(x) \log f(x) \le -\int_{-\infty}^{+\infty} f(x) \log g(x)\, dx$, при чему једнакост важи ако и само ако $f(x) = g(x)$ за свако реално x. База логаритма је произвољан реалан број $b > 1$.

Доказ: Полазимо од познате [14. Пример 1] неједнакости $\log t \le t - 1$, где знак једнакости вреди ако и само ако је $t = 1$. Нека је, за дато x: $f(x) > 0$ и $g(x) > 0$. Сменом $t = g/f$ добијамо:

$$\log \frac{g(x)}{f(x)} \le \frac{g(x)}{f(x)} - 1$$

$$f(x) \log g(x) - f(x) \log f(x) \le g(x) - g(x)$$

Ова неједнакост вреди и у тачки $f(x) = 0$, или $g(x) \to 0$. Непосредно интегрирање доказује став теореме. ■

Тврђење теореме следи и интуитивно, јер је интеграл $-\int f(x) \log f(x)\, dx$ инфинитезимални збир производа функције и логаритма функције, те (непрекидна) информација перцепције. Када је ова функција вредности мањих од један, онда је f растућа када је $-\log f$ опадајућа и обрнуто. А ако је функција већа од један, онда је она растућа када је $+\log f$ опадајућа и обрнуто. Дакле, оне су увек супротне монотоности, што значи минималну информацију перцепције. Она је зато мања од било којег другог интеграла производа f са логаритмом функције g и нарушавања ове монотоније.

Када је $0 < f < 1$ тада је $-\log f > 0$, а када је $f > 1$ тада је $+\log f > 0$. Зато променом предзнака ове теореме препознајемо и тврђење о максимуму информације перцепције, да се такав постиже када су фактори исте монотоности. Њен специјални случај је Шенонова информација непрекидне расподеле вероватноћа.

Пример 1. Нађимо Шенонову информацију Гаусове расподеле [36.], густине

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

средње вредности $\mu = 0$ и дисперзије σ. Логаритмовање даје $\ln \rho(x) = -\ln \sigma\sqrt{2\pi} - \frac{x^2}{2\sigma^2}$. Са друге стране, парцијалном интеграцијом налазимо:

$$1 = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} dx = \frac{x e^{-\frac{x^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \Big|_{-\infty}^{+\infty} + \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{x^2}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} dx$$

Први сабирак у обе граничне вредности је нула, а други, интеграл, са претходним логаритмом због норме даје

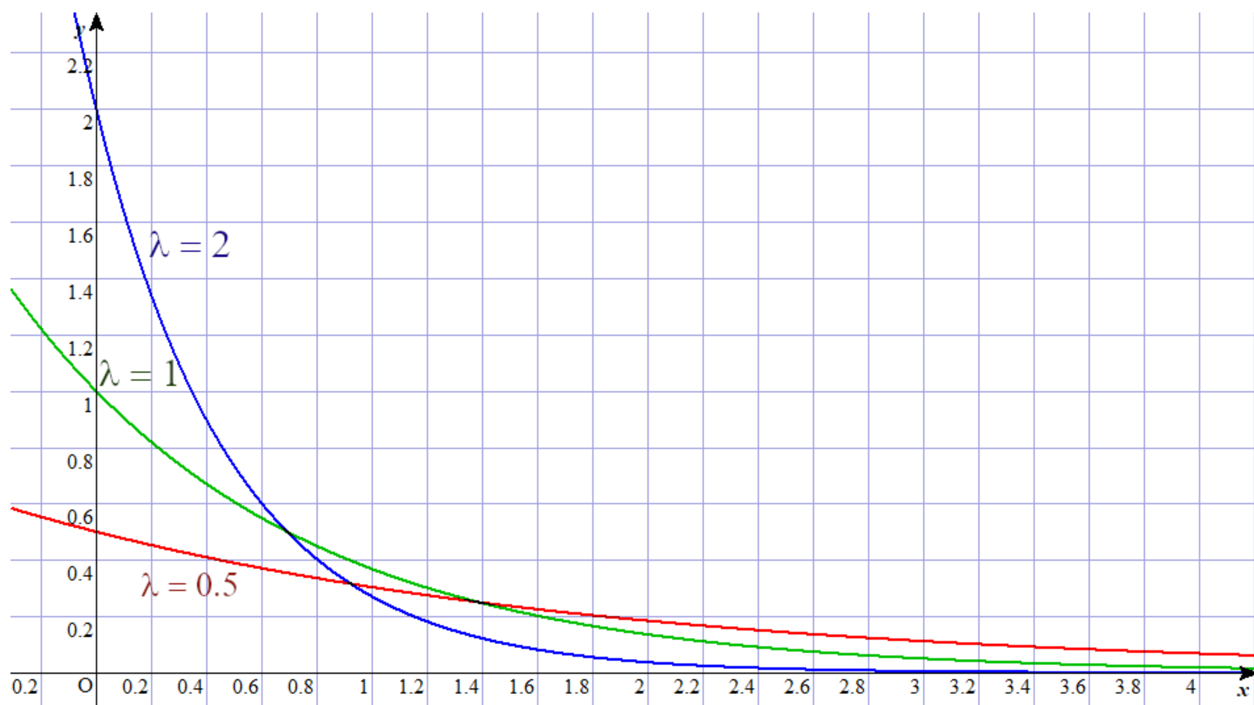
$$S = -\int_{-\infty}^{+\infty} \rho \ln \rho \, dx = \ln \sigma\sqrt{2\pi e}$$

и то је Шенонова информација Гаусове расподеле. Резултат је тачно једнак раније [36.] изведеном на мало другачији начин. ❑

Експоненцијална расподела случајне варијабле X дата је функцијом густине

$$\rho(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Где се константа $\lambda > 0$ назива „параметар брзине“. Она одређује средњу вредност ове расподеле, $E(X) = \frac{1}{\lambda} = \mu$, и варијансу $\mathrm{Var}(X) = \frac{1}{\lambda^2} = \sigma^2$. На први поглед она личи на Пуасонову густину, иако је потоња врста степене расподеле [34. Дефиниција 2. и 35.]. Слично се и примењује. На пример, трошење новца купаца у обиласку трговина прати експоненцијалну дистрибуцију. Више ће људи трошити мало, а мало их ће трошити велике количине новца. Таква је и расподела времена које службеник проводи са клијентом, или трајање дела компјутерске опреме до кварења и слично.



На графу видимо три функције густине експоненцијалне расподеле (редом за $\lambda = 0.5,\ 1,\ 2$). Што је већи λ оштрији је граф, средња вредност $\mu = \frac{1}{\lambda}$ ближа је нули, а дисперзија $\sigma = \frac{1}{\lambda}$ мања. Средња информација би, према томе, требала бити опадајућа функција параметра брзине λ.

Пример 2. Израчунавамо Шенонову информацију експоненцијалне расподеле:

$$S = -\int_0^{\infty} \rho \ln \rho \ dx = -\int_0^{\infty} \lambda e^{-\lambda x} \ln(\lambda e^{-\lambda x}) \ dx =$$

$$= \int_0^{-\infty} e^t \ln(\lambda e^t) \ dt = \int_0^{-\infty} e^t (\ln \lambda + t) \, dt$$

$$= \ln \lambda \int_0^{-\infty} e^t \ dt + \int_0^{-\infty} t e^t \ dt = (\ln \lambda) e^t \Big|_0^{-\infty} + \int_0^{-\infty} t \, de^t$$

$$= -\ln \lambda + \left(te^t - \int e^t \ dt \right)_0^{-\infty} = -\ln \lambda - \int_0^{-\infty} e^t \ dt$$

Према томе $S = 1 - \ln \lambda$. То је средња Хартлијева информација ове расподеле. ❑

Експоненцијална расподела је расподела вероватноће која се првенствено може користити када израчунавамо времена догађаја. Она је врста континуиране дистрибуције која помаже у процени временског трајања, односно времена када ће се одређени догађај вероватно догодити. Успешна је у праћењу много малих вредности са мало великих. У томе је слична степеној расподели. Важно својство експоненцијалне дистрибуције је непамћење, тј. одсуство утицаја бившег на вероватноћу будућег догађаја. Сада видимо и њен један проблем са Шеноновом информацијом.

Наиме, 2. пример дозвољава могућност, $S = 0$, да информација исчезне када је параметар брзине $\lambda = e = 2.71828 \ldots$, што би била немогуће мала вредност, упркос теоријски дозвољене вредности очекивања и дисперзије ове расподеле, $\mu = \sigma = 1/e$. То је био један од својевремених разлога за сумњу у Шенонову информацију и рад на књизи „Физичка информација[11]“, где сам показао да би информација за коју важи закон одржања морала бити мало већа. Навео сам тамо неке примере, предлоге таквих замена.

Иначе, нормална ($\ln \sigma\sqrt{2\pi e}$) и ова расподела ($\ln \sigma e$) имају веома сличне Шенонове информације, што нам што нам говори о њиховој заједничкој проблематици у континууму. Међутим, расподеле густина вероватноћа са оваквом дефиницијом информације имају проблеме и са „горње“ стране, у домену „блиском“ бесконачним вредностима случајних варијабли.

Пример 3. Следећа функција може бити густина расподеле вероватноћа

$$\varphi(x) = \begin{cases} \dfrac{1}{x \ln^2 x}, & x \geq e \\ 0, & x < e \end{cases}$$

јер је $\varphi(x) \geq 0$ за свако x, а уз то је и нормирана:

$$\int_{-\infty}^{+\infty} \varphi(x) \ dx = \int_{-\infty}^{e} \varphi(x) \ dx + \int_{e}^{+\infty} \varphi(x) \ dx =$$

[11] https://www.scribd.com/document/406574702/

$$= 0 + \int_{e}^{+\infty} \frac{dx}{x \ln^2 x} = \int_{e}^{+\infty} \frac{d \ln x}{\ln^2 x} = -\frac{1}{\ln x} \Big|_{e}^{\infty} = 1$$

Интеграл ове густине конвергира, али њена Шенонова информација дивергира:

$$S(\varphi) = -\int_{-\infty}^{+\infty} \varphi(x) \ln \varphi(x) \; dx = -\int_{e}^{+\infty} \frac{1}{x \ln^2 x} \ln \frac{1}{x \ln^2 x} \; dx =$$

$$= \int_{e}^{+\infty} \frac{\ln(x \ln^2 x)}{x \ln^2 x} \; dx = \int_{e}^{+\infty} \frac{\ln x + 2 \ln(\ln x)}{x \ln^2 x} \; dx$$

$$= \int_{e}^{+\infty} \frac{dx}{x \ln x} + 2 \int_{e}^{+\infty} \frac{\ln(\ln x)}{x \ln^2 x} \; dx \geq \int_{e}^{+\infty} \frac{dx}{x \ln x} = \ln(\ln x) \Big|_{e}^{+\infty} = +\infty$$

Коначна густина вероватноћа има бесконачну Шенонову информацију. ❏

Пример 4. У наслову о метричким просторима, овде [19. Пример 4], поменут је простор l_p низова $x = [\![\xi_k]\!]$, таквих да ред, норма вектора $\|x\| = \sum_{k=1}^{\infty} |\xi_k|^p$ конвергира. У l_p растојање је уведено са

$$d(x, y) = \left(\sum_{k=1}^{\infty} |\xi_k - \zeta_k|^p \right)^{1/p}$$

а под l подразумевамо l_1. Сличан је простор $L_p(a, b)$ функција $x(t)$ са $\int_a^b |x|^p \, dt < +\infty$ норме

$$\|x\| = \left(\int_{a}^{b} |x(t)|^p \, dt \right)^{1/p}$$

са параметром $p \geq 1$ и границама интервала $a < b$. ❏

Пример 5. Нека је $y(t)$ произвољна фиксирана функција из простора $L_p(a, b)$. Тада из Хелдерове неједначине следи $|x^*(x)| \leq \left(\int_a^b |y|^p \;\; dt \right)^{\frac{1}{p}} \left(\int_a^b |x|^q \;\; dt \right)^{\frac{1}{q}} = \|y\|_{L_p} \|x\|_{L_q}$, $x \in L_q$, $\frac{1}{p} + \frac{1}{q} = 1$, те је линеарна функционела на том простору $x^*(t) = \int_a^b y(t) x(t) \, dt$ и $\|x^*\| \leq \|y\|_{L_p}$ ограничена. ❏

Овакви примери функционалне анализе лако се преносе на „информацију прерцепције“, ради дубљег разумевања и генералисања Шенонове информације.

41. Ограничења

Видели смо [40. Пример 1] да нормална расподела, густине $\rho_N(x) = \exp(-x^2/2\sigma^2)/(\sqrt{2\pi}\sigma)$, има Шенонову информацију $S_N(\sigma) = \ln(\sqrt{2\pi e}\sigma)$. Затим [40. Пример 3] да експоненцијална расподела, густине $\rho_E(x) = \frac{1}{\sigma}\exp\left(-\frac{x}{\sigma}\right)$ за $x \geq 0$ и $\rho_E = 0$ за $x < 0$, има Шенонову информацију $S_E = \ln(e\sigma)$.

Пример 1. Униформна (равномерна) расподела на интервалу (a,b) задата је густином

$$\rho_U(x) = \begin{cases} \dfrac{1}{b-a}, & x \in (a,b) \\ 0, & x \notin (a,b) \end{cases}$$

Интеграл густине графички представља површину између функције густине и апсцисе, која је због нормираности расподеле, увек јединична. Шенонова информација ове расподеле је:

$$S_U = -\int_{-\infty}^{+\infty} \rho_U(x) \ln \rho_U(x)\ dx = \int_a^b \frac{\ln(b-a)}{b-a}\ dx = b-a$$

Приметимо да ова информација, зависно од дужине интервала, може бити сваки реалан број па и негативна. ❑

У сва три случаја следеће теореме, која говори о Шеноновој информацији континуума, користимо претходну теорему [40. Теорема 1].

Теорема 1. Нека је дата густина вероватноће $f(x)$ са непрекидном реалном променљивом x.

i. Ако је $f(x) = 0$ свугде изван интервала $(a,b) \subseteq [0,1]$, тада постоји информација за коју важи $S \leq \log(b-a)$, при чему вреди једнакост ако и само ако је f униформна на (a,b).

ii. Ако је $f(x) = 0$ за $x < 0$ и ако постоји коначно очекивање μ тада постоји информација и важи $S \leq \log(e\mu)$, при чему једнакост вреди ако и само ако f има експоненцијалну расподелу.

iii. Ако постоји коначна варијанса σ^2, тада постоји информација таква да је $S \leq \log(\sqrt{2\pi e}\sigma)$, при чему једнакост вреди ако и само ако f има нормалну расподелу.

Доказ: *i*. Узмимо за g униформну расподелу из 1. примера. Тада је:

$$S = -\int_{-\infty}^{+\infty} f(x) \ln f(x)\ dx \leq -\int_{-\infty}^{+\infty} f(x) \ln g(x)\ dx =$$

$$= -\int_{-\infty}^{+\infty} f(x)\,[-\ln(b-a)]\,dx = \ln(b-a)$$

Последња једнакост долази од нормираности f. Тиме је доказан први део теореме (*i*).

ii. Узмимо за g експоненцијалну расподелу [40.]. Тада:

$$S = -\int_0^{+\infty} f(x) \ln f(x) \; dx \leq -\int_0^{+\infty} f(x) \ln g(x) \; dx = -\int_0^{+\infty} f(x) \left(-\ln \mu - \frac{x}{\mu} \right) \, dx =$$

$$= \ln \mu + \frac{1}{\mu} \int_0^{+\infty} x \, f(x) \, dx = \ln \mu + 1 = \ln \mu + \ln e = \ln(e\mu)$$

Множећи са константом $\log_b e$ прелазимо на произвољну базу логаритма. Тиме је доказано (*ii*).

iii. Узмимо за g нормалну расподелу, такође из претходног наслова:

$$S = -\int_0^{+\infty} f(x) \ln f(x) \; dx \leq -\int_{-\infty}^{+\infty} f(x) \ln g(x) \; dx = -\int_{-\infty}^{+\infty} f(x) \left(-\ln \sigma\sqrt{2\pi} - \frac{x^2}{2\sigma^2} \right) \, dx =$$

$$= \ln \sigma\sqrt{2\pi} + \frac{1}{2\sigma^2} \int_{-\infty}^{+\infty} x \, f(x) \, dx = \ln \sigma\sqrt{2\pi} + 0.5 = \ln \sigma\sqrt{2\pi e}$$

Тиме је доказан и трећи део (*iii*) теореме. ■

Са доње стране ограничења је свет инфинитезимала. Претпостављамо да је x-оса подељена на мале интервале $\Delta x_k > 0$ са вероватноћама $P(\Delta x_k) \geq 0$ које називамо елементима вероватноће. Јасно је да за елементе вероватноће важи $\sum_k P(\Delta x_k) = 1$. Даље претпоставимо да су тачке унутар интервала јернако вероватне, у том смислу да можемо дефинисати $p(x) = P(\Delta x_k)$, за $x \in \Delta x_k$.

Вероватноће различитих интервала могу бити различите, али тачака унутар њих не. За појашњење густине вероватноће дефинишемо Δx „стандардну дужину“. Та основна дужина интервала је n-ти део јединичне дужи, $\Delta x = 1/n$, где је n неки природан број. Густина вероватноће постаје гранична вредност количника елемената вероватноће и стандардне дужине

$$\rho(x) = \lim_{\Delta x \to 0} \frac{p(x)}{\Delta x}$$

где је произвољно бирано $x \in \Delta x$. Ова „густина вероватноће“ аналогна је „густини супстанце“, у физици $\gamma = m/\Delta V$, која је количник масе m и основне запремине ΔV коју та маса заузима.

Однос две густине вероватноћа у околини тачке x је однос њихових вероватноћа интервала исте дужине, $\rho_1(x) : \rho_2(x) = p_1(x) : p_2(x)$. Ставимо ли да је други од интервала основни са униформном расподелом, $\rho_2 = 1$, када је $p_2 = \Delta x$, ова пропорција даје $p(x) = \rho(x) \, \Delta x$, $\Delta x \to 0$. Сабирајући све ове вероватноће, узимајући само по једну тачку из појединог интервала, у граничном случају биће

$$\int_{-\infty}^{+\infty} \rho(x) \, dx = 1$$

Нормираност густина расподеле вероватноћа аналогна је нормираности расподеле дискретних вероватноћа на начин који нам одговара.

Густина није буквално вероватноћа и потребно нам је њено усклађивање са првим разматрањима. Доследно, настају разлике у интерпретацији основних појмова теорије информације, од којих сам неке већ помињао уз Шенонову информацију. Даље разлике потражимо путем логаритмовања дефиниције густине

$$-\log\rho = \lim_{\Delta x\to 0}(-\log p + \log\Delta x)$$

Ту је „$-\log p$“ информација датог интервала, затим „$-\log\Delta x$“ информација основног интервала, а „$-\log\rho$“ је разлика претходне две информације. Средња вредност (математичко очекивање) ове густине је

$$S^*(X) = -\int_{-\infty}^{+\infty}\rho(x)\log\rho(x)\ dx$$

и то је „информација континуума“.

Дискретна информација S је математичко очекивање (средња вредност) информације елемената дискретног скупа. Информација континуума S^* је математичко очекивање разлика информација елемената дате и основне расподеле континуума. Погледајмо то детаљније.

Пример 2. Средња вредност $p(x) = \rho(x)\,\Delta x$, када $\Delta x \to 0$, даје дискретну информацију:

$$S(x) = -\sum_x [\rho(x)\,\Delta x]\log[\rho(x)\Delta x] =$$

$$= -\sum_x [\rho(x)\log\rho(x)]\,\Delta x - \sum_x \rho(x)\,\Delta x\log\Delta x$$

где узимамо по једно x из сваког од интервала. Пустимо ли да Δx тежи нули, први збир постаје претходни интеграл, а други постаје интеграл пре тога, тј.

$$S(x) = S^*(x) - \log\Delta x$$

када $\Delta x \to 0$. Та је релација писана недовршено, јер $\log\Delta x \to \infty$, али нам у овом облику показује да дискретну информацију можемо раставити на два сабирка, од којих је први информација континуума (углавном) коначан број, а формално јако подсећа на дискретну информацију. ❑

То је (један од) прелаз са дискретне на континуалну информацију, са чувањем интуитивне везе међу основним појмовима. Укатко, из претходног следи:

$$\begin{cases} S^* = \lim_{\Delta x\to 0}(S + \log\Delta x), \\ \lim_{\Delta x\to 0}\dfrac{S}{\log\Delta x} = -1. \end{cases}$$

Из прве једнакости следи да информација континуума не зависи од Δx, па је можемо писати у облику $\lim_{\Delta x\to 0}\left(\frac{S}{\log\Delta x} + 1 - \frac{S^*}{\log\Delta x}\right) = 0$ где трећи сабирак тежи нули и отуда друга једнакост. Основна дужина има инфинитезималну информацију истог реда са дискретном и континуалном.

42. Пренос варијабле

Разматрања ширимо на 2-димензионалну непрекидну расподелу, држећи се постојеће теорије вероватноће. Случајну промењљиву не посматрамо само као прост 2-дим скуп неодређености, већ као пренос са једне 1-дим на другу 1-дим неодређеност континуума.

Каналом $K: X \to Y$ преноси се тачка $x \in X$ у тачку $y \in Y$ са вероватноћом $p(x, y)$. Када су X и Y дискретни скупови, тада се канал може описати матрицом $\widehat{K}$ како смо то радили у првом делу ове књиге, али ако је бар један од та два скупа непребројив суочићемо се са нула вероватноћама и третираћемо их на начин обрађен у претхоном наслову.

Претпоставимо да су улазни (X) и излазни (Y) догађаји непребројиви скупови тачака апсцисе и ординате. Поделимо ове осе на интервале дужина Δx и Δy, тако да добијамо пребројиве скупове интервала и дискетно пресликавање вероватноћа. Ако уочимо један пар интервала $(\Delta x, \Delta y)$ могуће је говорити о вероватноћи $P(\Delta x, \Delta y)$ да скуп тачака из Δx буде пресликан унутар Δy.

Нека су ови интервали веома малих дужина и да постоји лимес

$$k(x, y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P(\Delta x, \Delta y)}{\Delta x \cdot \Delta y}$$

где $x \in \Delta x$ и $y \in \Delta y$. Функција $k(x, y)$, ако постоји, представља густину расподела вероватноћа канала K. Подразумева се да су интервали међусобно дисјунктни и да њихова унија даје X и Y осу.

Збир свих вероватноћа интервала дуж y-осе (односно x-осе), у које се преноси дати интервал Δx (Δy), представља вероватноћу реализације интервала Δx (Δy):

$$p(\Delta x) = \sum_{y \in \Delta y} P(\Delta x, \Delta y)$$

$$q(\Delta y) = \sum_{x \in \Delta x} P(\Delta x, \Delta y)$$

У збиру узимамо по једну тачку унутар сваког од интервала Δy (односно Δy). Добија се дискретна маргинална расподела x-осе (y-осе). Сменом претходне једнакости овде, добијамо:

$$p(\Delta x) = \sum_{y \in \Delta y} k(x, y)\, \Delta x \Delta y$$

$$q(\Delta y) = \sum_{x \in \Delta x} k(x, y)\, \Delta x \Delta y$$

а узмемо ли у обзир претходну дефиницију густине, преласком на лимес биће

$$\rho_y(x) = \int_{-\infty}^{+\infty} k(x, y)\, dy$$

$$\rho_x(y) = \int_{-\infty}^{+\infty} k(x,y)\,dx$$

Ако је $k(x,y)$ интеграбилна функција, тада постоје ове густине, непрекидна маргинална густина расподеле апсцисе (x-осе), односно ординате (y-осе).

Пођемо ли опет од исте горње вероватноће, због нормираности интервала, имамо

$$\sum_x \sum_y P(\Delta x, \Delta y) = 1$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k(x,y)\,dxdy = 1$$

ове интеграле добијамо сменом густине и преласком на лимес. Нормираност густине $k(x,y)$ сада повлачи нормираност густина ρ:

$$\int_{-\infty}^{+\infty} \rho_y(x) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k(x,y)\,dy = 1$$

и аналогно $\int_{-\infty}^{+\infty} \rho_x(y) = 1$. Са друге стране, због ненегативности вероватноћа, густине k, ρ_y и ρ_x такође су ненегативне за свако x и y.

Пример 1. (Црна кутија) Нека је канал представљен густином који се може исказати производом две 1-дим густине, $k(x,y) = f_1(x) \cdot f_2(y)$. Тада су према претходном, густине расподеле улазних односно излазних сигнала датог канала $\rho_y(x) = f_1(x)$ и $\rho_x(y) = f_2(y)$. Излазна информација, коју „преноси“ овај канал ($k: x \to y$) не зависи од послате, али и обрнуто ($k^{-1}: y \to x$) послата не зависи од примљене. Такав канал не преноси информацију уопште. ❑

Можда и занимљивије од „Црне кутије“ било би растављање густине вероватноће на сабирке. Чак и са само три такве компоненте, $k(x,y) = k_1(x,y) + k_2(x,y) + k_3(x,y)$, густина вероватноће била би прешироко тема за овај концепт „теорије информације“. На пример, квадратна форма, која је полином другог степена од n променљивих, може се писати у облику:

$$k(x,y) = af_1^2(x) + bf_1(x)f_2(y) + cf_2^2(y) = (f_1 \quad f_2)\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \vec{f}^{\tau}\hat{A}\vec{f}$$

где се разматрање континуалних функција, $f_1 = f_1(x)$ и $f_2 = f_2(y)$, даље је и дискретни задатак који се обрађује рецимом начинима описаним у првом делу ове скрипте.

43. Условне вероватноће

Вероватноћа да променљива x прими вредност из интервала $(x, x + dx)$, без обзира коју вредност при томе прими променљива y, једнака је елементу вероватноће $\rho_y(x)dx$. Зато је вероватноћа да y буде у интервалу $(y, y + dy)$, ако је x било у интервалу $(x, x + dx)$, једнака

$$k(y|x)dy = \frac{k(x,y)dy}{\rho_y(x)}$$

Обратно, вероватноћа да x падне у интервал $(x, x + dx)$, када је извесно $y \in (y, y + dy)$, је

$$k(x|y)dx = \frac{k(x,y)dx}{\rho_y(x)}$$

Ове две функције представљају густине расподела условних вероватноћа. Следећи пример сличан је недавни [41. Пример 2].

Пример 1. Полазећи од дискретне (Шенонове) информације канала по интервалима:

$$S(X,Y) = -\sum_x \sum_y k(x,y)\,\Delta x \Delta y \cdot \log[k(x,y)\Delta x \Delta y] =$$

$$= \sum_x \sum_y k(x,y)\,\Delta x \Delta y \cdot \log k(x,y) - \log(\Delta x \Delta y)$$

У последњем сабирку је узета у обзир нормираност расподеле $k(x,y)$. Прелазећи на граничне вредности, $\Delta x \to 0$ и $\Delta y \to 0$, први збир постаје информација непрекидног канала

$$S^*(X,Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k(x,y) \log k(x,y)\, dxdy$$

При томе је $S(X,Y) = S^*(X,Y) - \log(\Delta x \Delta y)$, где $\Delta x \to 0$, $\Delta y \to 0$. ❑

Разлика дискретне и непрекидне информације канала расте са смањивањем стандардних дужина, тј. $S - S^* = -\log(\Delta x \Delta y) \to +\infty$. Другим речима, како је непрекидна информација канала коначна, $S^* = S + \log(\Delta x \Delta y)$, следи да је дискретна S истог реда величине као и (Хартлијева) информација „стандардне површине“, $-\log(\Delta x \Delta y)$.

Дискретна условна (Шенонова) информација X-осе за дато Δy је

$$S(X|\Delta y) = -\sum_x k(x|\Delta y)\,\Delta x \cdot \log[k(x|\Delta y)\Delta x]$$

Средња вредност ових бројева дуж Y-осе је:

$$S(X|Y) = \sum_y \rho_x(y)\,\Delta y \cdot S(X|\Delta y) =$$

$$= -\sum_y \sum_x k(x,y)\,\Delta x\,\Delta y \cdot \log k(x|y) - \log \Delta x$$

Преласком на лимес, први збир постаје постаје условна непрекидна информација

$$S^* = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} k(x,y)\log k(x|y)\,dxdy$$

Приметимо да су дискретна и непрекидна ова информација такође у релацији

$$S(X|Y) = S^*(X|Y) - \log \Delta x$$

када $\Delta x \to 0$.

Слично, полазећи од дискретне условне информације Y-осе, преласком на лимес добијамо

$$S^*(Y|X) = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} k(x,y)\log k(y|x)\,dxdy$$

Овде имамо релацију

$$S(Y|X) = S^*(Y|X) - \log \Delta y$$

када $\Delta y \to 0$. За разлику од информације канала, горе поменуте $S(X,Y)$, ова условна информација $S(Y|X)$ дефинише „шум" ($X \to Y$) канала. Понавља се да се и дискретни шум од непрекидног шума разликује за (Хартлијеву) информацију стандардне дужине Δy, или инфинитезималног дела Y-осе.

Ове исте разлике излазне информације дискретних интервала $S(Y)$, а са друге стране непрекидне информације $S^*(Y)$ горе поменуте, су

$$S(Y) - S(Y|X) = S^*(Y) - S^*(Y|X)$$

и инваријантне су на $\Delta y \to 0$. Оне иду у прилог изложеног преласка са дискретне информације на непрекидну.

Држимо ли и даље да је $S^*(Y|X)$ шум непрекидног канала, а узајамна информација непрекидног канала $S^*(Y,X) = S^*(Y) - S^*(Y|X)$, тада информацију излаза можемо у потпуности раставити на збир узајамне информације и шума континуума

$$S^*(XY) + S^*(Y|X) = S^*(Y)$$

При томе су шум дискретног (S) и шум непрекидног (S^*) канала ($\Delta y \to 0$) у релацији

$$S(X,Y) = S^*(X,Y) - \log \Delta y$$

док је узајамна информација према претходном инваријантна, $S(XY) = S^*(XY)$. Погледајмо ово на поменутој „црној кутији" [42. Пример 1].

Пример 2. Канал има густину $k(x,y) = f_1(x) \cdot f_2(y)$. Густине условних расподела су $k(y|x) = f_2(y)$ и $k(x|y) = f_1(x)$, јер промењљива y не зависи од x, нити обрнуто. Другим речима, из примљеног сигнала y не можемо сазнати било шта о улазном сигналу, нити обрнуто, узајамне информације нема, $S^*(XY) = 0$, док је шум канала једнак информацији излаза, $S^*(Y|X) = S^*(Y)$.

Заиста, полазећи од претходних формула:

$$S^*(Y|X) = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} k(x,y)\log k(y|x)\,dxdy =$$

$$= -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_1(x)\,f_2(y)\log f_2(y)\,dxdy$$

$$= -\int_{-\infty}^{+\infty} f_1(x)\,dx \cdot \int_{-\infty}^{+\infty} f_2(y)\log f_2(y)\,dy = S^*(y)$$

Даље је очигледно $S^*(X,Y) = 0$. ❑

Горе смо видели да је узајамна информација инваријанта дискретног и непрекидног канала. Другу врсту инваријанти дају међусобни односи неких дискретних информација, као на пример:

$$\begin{cases} S(X,Y) = S(Y) + S(X|Y) \\ S(X,Y) = S(X) + S(Y|X) \end{cases}$$

а отуда и одговарајући односи непрекидних информација:

$$\begin{cases} S^*(X,Y) = S^*(Y) + S^*(X|Y) \\ S^*(X,Y) = S^*(X) + S^*(Y|X) \end{cases}$$

Једнакости се добијају крађењем за $\log \Delta x$ и $\log \Delta y$, али исте се, наравно, могу доказати и интегрирањем.

Пример 3. Густина канала је дата нормалном расподелом

$$k(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x}{\sigma_x}\right)^2 - 2\rho\frac{x}{\sigma_x}\frac{y}{\sigma_y} + \left(\frac{y}{\sigma_y}\right)^2\right]\right\}$$

Густина расподеле улазних, односно излазних сигнала, у ознакама [42.], су у теорији вероватноће познате маргиналне расподеле:

$$\rho_y(x) = \frac{1}{\sigma_x\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x}{\sigma_x}\right)^2\right\}, \quad \rho_x(y) = \frac{1}{\sigma_y\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{y}{\sigma_y}\right)^2\right\}$$

Познато је такође да је за сваки фиксиран x променљива y нормално дистрибуирана и то са константном (условном) дисперзијом

$$\sigma_{y|x} = \sigma_y\sqrt{1-\rho^2}$$

и очекивањем које зависи линеарно од x-а:

$$\mu_{y|x} = \rho \cdot \frac{\sigma_y}{\sigma_x}$$

$$k(y|x) = \frac{1}{\sigma_{y|x}\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{y-\mu_{y|x}}{\sigma_{y|x}}\right)^2\right\}$$

Приметимо да коефицијент корелације ($0 \le |\rho| \le 1$) одређује зависност излаза (y) од улаза (x). Када је $\rho = 0$ имамо $\sigma_{y|x} = \sigma_y$ и $\mu_{y|x} = 0$, па је $k(y|x) = \rho_x(y)$, тј. излаз не зависи од улаза, та је дати канал „црна кутија“. Када $|\rho| \to 1$ променљиве x и y су у линеарној функционалној вези и имамо „белу кутију“. У осталим случајевима (0<$|\rho| < 1$) израчунавамо у натима:

$$S^*(X,Y) = \ln\left(\sigma_x \sigma_{y|x} 2\pi e\right)$$

$$S^*(Y) = \ln\left(\sigma_y \sqrt{2\pi e}\right)$$

$$S^*(Y|X) = \ln\left(\sigma_{y|x} \sqrt{2\pi e}\right)$$

$$S^*(XY) = \ln\left(\sigma_y / \sigma_{y|x}\right)$$

У сваком случају збир средње узајамне информације и шума даје излазну информацију непрекидног канала:

$$S^*(XY) + S^*(Y|X) = S^*(Y)$$

при чему је $0 \le S^*(XY) < +\infty$, $-\infty < S^*(Y|X) < +\infty$ и $0 \le \sigma_{y|x} \le \sigma_y$. Када је $\sigma_{y|x} = \sigma_y$ тада $S^* = 0$ и канал је непропустан (црна кутија). Када $\sigma_{y|x} \to 0$ имамо $S^* \to +\infty$ и $S^*(Y|X) \to -\infty$. Отуда следи да је „капацитет“ непрекидног канала неограничен и да је бела кутија непрекидни канал са минус бесконачно великим шумом. ❑

Овим информацијама треба додати информацију перцепције, која их делом и поопштава, а која у свом основном облику говори о спрези $S = \vec{a} \cdot \vec{b}$ два низа, $\vec{a} = (a_1, a_2, \dots, a_n)$ и $\vec{b} = (b_1, b_2, \dots, b_n)$, субјектових способности и објективних околности. Ма која метрика [19.] простора вероватноћа да је у питању, могуће је дефинисати нормирану, јединичну информацију перцепције

$$S_{||} = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$$

која је ненегативан број не већи од један, $0 \le S_{||} \le 1$. Представљен у равни паралелограма који разапињу дати низови, ово је косинус угла између њих, $S_{||} = \cos \sphericalangle(\vec{a}, \vec{b})$.

Пример 4. У 3-Д простору вероватноћа [38. Пример 1] вероватнији исходи групишу се унутар лопте полупречника R, са тачком на њеној сфери чији радијус дејства је r. Мањи полупречник сматрамо занемарљивим у односу на велики и налазимо однос заједничке запремине мање и веће лопте са преосталим делом мање лопте $V_1 : V_2 = 1 - l/R$, где је $l = 3r/4$ константно. Овај однос тумачили смо „потенцијалом“ који унутрашњост веће лопте држи даље од вањске неизвесности.

Овај је потенцијал, као и нормирана информација перцепције, број између 0 и 1. Зато можемо писати $1 - \frac{q}{R} = \varepsilon S_{||}$ где је ε нека константа која изједначава физичке димензије. Отуда, узимајући извод по времену и користећи горњи косинус, редом добијамо једнакости:

$$\frac{q}{R^2} = \varepsilon \frac{\partial S_{||}}{\partial t}, \quad R = \frac{l}{1 + \varepsilon \cos \sphericalangle(\vec{a}, \vec{b})}$$

Прва говори о „силама“, а друга о коникама по којима би оне покретале „набој“. ❑

Прилози

У програмском језику Python 3.9 радио сам модуле са функцијама током тестирања теза у тексту, од којих сам неке сачувао. Прва функција (povrs1) генерише случајни вектор x и њему очекивање вектор z, за дати број M понављања опита, и израчунава површину П паралелограма који вектори разапињу. Њу неколико пута позива следећа функција (linea1) која тако формира парове (M, П) и за њих израчунава линеарну регресију (y = a + bx).

Слично се понавља са следеће две функције, прва (povrs2) која за M понављања опита генерише вектор x и за корак више (за dM) генерише суседни случајни вектор y, те израчунава површину П коју они разапињу. Наредна функција (linea2) користи ову да формира низ парова (M, П) и нађе линеарну процену (y = a + bx) као регресију парова случајних варијабли. Све четири ове функције служиле су ми за израчунавања примера регресије [23-24 и 27.] .

```
# D:/Python/reg - modul sa funkcijama slucajnih varijabli

import sys
sys.path.append('D:\Python')
#sys.path.insert(0, "D:/Python")
import reg

def povrs1(br1=100, br2=1):
    # Unesi dva broja br1 i br2, inace br1 = 100 i br2 = 1
    M = br1
    import random
    import math
    x1 = x2 = x3 = x4 = 0
    #verovatnoce
    p1 = 0.27 # 0.6 # 0.4
    p2 = 0.26 # 0.2 # 0.3
    p3 = 0.24 # 0.1 #0.2
    p4 = 0.23 # 0.1
    for m in range(M):
        r = random.random()
        if r <= p1: x1 = x1 + 1
        if r > p1 and r <= p1 + p2: x2 = x2 + 1
        if r > p1 + p2 and r <= p1 + p2 + p3: x3 = x3 + 1
        if r > p1 + p2 + p3 and r <= p1 + p2 + p3 + p4: x4 = x4 + 1
    m1 = M*p1
    m2 = M*p2
    m3 = M*p3
    m4 = M*p4
    stampa1 = "M = %2d, x = (%2d, %2d, %2d, %2d), z = (%1.1f, %1.1f,
%1.1f, %1.1f)" % (M, x1, x2, x3, x4, m1, m2, m3, m4)
    xz = x1*m1 + x2*m2 + x3*m3 + x4*m4
    x0 = math.sqrt(x1**2 + x2**2 + x3**2 + x4**2)
    z0 = math.sqrt(m1**2 + m2**2 + m3**2 + m4**2)
    stampa2 = "x.z = %1.2f, |x| = %1.2f, |z| = %1.2f" % (xz, x0, z0)
    kosin = xz/(x0*z0)
```

```
    sinus = math.sqrt(1 - kosin**2)
    # povrs = x0*z0*sinus
    povrs = math.sqrt(x0*z0)*sinus # 1. zamena
    stampa3 = "cos fi = %1.5f, sin fi = %1.5f, povrs = %1.2f" %
(kosin, sinus, povrs)
    if br2 == 1:
        print(stampa1)
        print(stampa2)
        print(stampa3)
    return povrs


def linea1(br1=100, br2=1):
    x1 = br1 # Uneti startni broj br1, inace br1 = 100. Drugi broj,
ako je 1 dodatna stampanja
    dx = 10 # Korak uvecanja broja
    k = 4 # Broj koraka
    y1 = povrs1(x1,br2)
    x2 = x1 + dx
    y2 = povrs1(x2,br2)
    x3 = x2 + dx
    y3 = povrs1(x3,br2)
    x4 = x3 + dx
    y4 = povrs1(x4,br2)
    sx = x1 + x2 + x3 + x4
    sy = y1 + y2 + y3 + y4
    sx2 = x1**2 + x2**2 + x3**2 + x4**2
    sxy = x1*y1 + x2*y2 + x3*y3 + x4*y4
    brojnik1 = sy*sx2 - sx*sxy
    brojnik2 = k*sxy - sx*sy
    nazivnik = k*sx2 - sx**2
    a = brojnik1/nazivnik
    b = brojnik2/nazivnik
    stampa1 = "sum_x = %1.2f, sum_y = %1.2f, sum_x2 = %1.2f, sum_xy =
%1.2f" % (sx, sy, sx2, sxy)
    stampa2 = "a = %1.5f, b = %1.5f" % (a, b)
    if br2 == 1:
        print(stampa1)
        print(stampa2)
    print("regresija: y = a + bx")
    return a, b


def povrs2(br1=10000, br2=1):
    # Unesi dva broja br1 i br2, inace br1 = 10000 i br2 = 1
    Mx = br1
    My = Mx + 1000
```

```
    import random
    import math
    x1 = x2 = x3 = x4 = 0
    y1 = y2 = y3 = y4 = 0
    #verovatnoce
    p1 = 0.4 # 0.6 # 0.27
    p2 = 0.3 # 0.2 # 0.26
    p3 = 0.2 # 0.1 # 0.24
    p4 = 0.1 # 0.1 # 0.23
    for m in range(Mx):
        r = random.random()
        if r <= p1: x1 = x1 + 1
        if r > p1 and r <= p1 + p2: x2 = x2 + 1
        if r > p1 + p2 and r <= p1 + p2 + p3: x3 = x3 + 1
        if r > p1 + p2 + p3 and r <= p1 + p2 + p3 + p4: x4 = x4 + 1
    for m in range(My):
        r = random.random()
        if r <= p1: y1 = y1 + 1
        if r > p1 and r <= p1 + p2: y2 = y2 + 1
        if r > p1 + p2 and r <= p1 + p2 + p3: y3 = y3 + 1
        if r > p1 + p2 + p3 and r <= p1 + p2 + p3 + p4: y4 = y4 + 1
    stampa1 = "Mx = %2d, x = (%2d, %2d, %2d, %2d), My = %2d, y = (%2d,
%2d, %2d, %2d)" % (Mx, x1, x2, x3, x4, My, y1, y2, y3, y4)
    xy = x1*y1 + x2*y2 + x3*y3 + x4*y4
    x0 = math.sqrt(x1**2 + x2**2 + x3**2 + x4**2)
    y0 = math.sqrt(y1**2 + y2**2 + y3**2 + y4**2)
    stampa2 = "x.y = %2d, |x| = %1.2f, |y| = %1.2f" % (xy, x0, y0)
    kosin = xy/(x0*y0)
    sinus = math.sqrt(1 - kosin**2)
    # povrs = x0*y0*sinus
    povrs = math.sqrt(x0*y0)*sinus # 1. zamena
    stampa3 = "cos fi = %1.5f, sin fi = %1.5f, povrs = %1.2f" %
(kosin, sinus, povrs)
    if br2 == 1:
        print(stampa1)
        print(stampa2)
        print(stampa3)
    return povrs


def linea2(br1=10000, br2=1):
    x1 = br1 # Uneti startni broj br1, inace br1 = 10000. Drugi broj,
ako je 1 dodatna stampanja
    dx = 1000  # Korak uvecanja broja
    k = 4 # Broj koraka
    pov1 = povrs2(x1,br2)
    x2 = x1 + dx
```

```
    pov2 = povrs2(x2,br2)
    x3 = x2 + dx
    pov3 = povrs2(x3,br2)
    x4 = x3 + dx
    pov4 = povrs2(x4,br2)
    sx = x1 + x2 + x3 + x4
    sy = pov1 + pov2 + pov3 + pov4
    sx2 = x1**2 + x2**2 + x3**2 + x4**2
    sxy = x1*pov1 + x2*pov2 + x3*pov3 + x4*pov4
    brojnik1 = sy*sx2 - sx*sxy
    brojnik2 = k*sxy - sx*sy
    nazivnik = k*sx2 - sx**2
    a = brojnik1/nazivnik
    b = brojnik2/nazivnik
    stampa1 = "sum_x = %1.2f, sum_y = %1.2f, sum_x2 = %1.2f, sum_xy = 
%1.2f" % (sx, sy, sx2, sxy)
    stampa2 = "a = %1.5f, b = %1.5f" % (a, b)
    if br2 == 1:
        print(stampa1)
        print(stampa2)
    print("regresija: y = a + bx")
    return a, b
```